INTERIM REPORT

Analysis of Full-Coverage Magnetometer Data in Relation to Statistically-Based Site Characterization Tools

Pueblo Precision Bombing and Pattern Gunnery Range #2
ESTCP Wide Area Assessment Demonstration

ESTCP Project MM-0325

September 2008

Dr. Brent Pulsipher
Pacific Northwest National Laboratory

Environmental Security Technology Certification Program

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
SEP 2008		**00-00-2008 to 00-00-2008**

4. TITLE AND SUBTITLE	5a. CONTRACT NUMBER
Analysis of Full-Coverage Magnetometer Data in Relation to Statistically-Based Site Characterization Tools: Pueblo Precision Bombing and Pattern Gunnery Range #2 ESTCP Wide Area Assessment Demonstration	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Pacific Northwest National Laboratory,902 Battelle Boulevard,Richland,WA,99354**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**	**Same as Report (SAR)**	**59**	

Contents

List of Figures

List of Tables

Executive Summary

The recently completed Wide Area Assessment (WAA) at the Pueblo Precision Bombing and Pattern Gunnery Range (PPBR) provides a unique opportunity to compare geophysical data sets collected from both aerial and ground-based platforms, as well as use areas of 100 percent survey coverage to verify statistical approaches to UXO site characterization. Four distinct data sets have been collected at the PPBR: 1) A "sparse" ground-based magnetometer dataset collected along geophysical transects that cover 0.64 percent of the site area; 2) A "dense" ground-based transect dataset that covers 1.64 percent of the total site area and includes the sparse transect data set; 3) Ten small regions of the site within which 100 percent survey coverage was achieved using the ground-based magnetometer sensors; and 4) 100 percent coverage of the entire site with helicopter mounted magnetometers.

The main issues addressed in this report are: 1) The consistency of the ground-based and helicopter-based magnetometer surveys in the areas where 100 percent coverage is available for both sensor modes; 2) The ability of the geostatistical kriging estimator to estimate spatially varying anomaly density from limited transect sampling as defined by comparison with the 100 percent helicopter-based surveys; 3) The ability of the kriging estimator to accurately estimate spatially varying densities and delineate high anomaly density areas within the 10 regions of 100 percent ground-based survey coverage from only limited transect data; and 4) The sensitivity of transect-based identification and delineation of high anomaly density areas to variations in the origin and orientation of the transects.

Direct comparison of the number of geophysical anomalies identified by ground-based and helicopter-based sensor platforms in areas of joint 100 percent coverage show that the helicopter mounted sensors identify only 8 to 29 percent of the anomalies identified by the ground-based sensors. The variation in the ability of the helicopter mounted sensors to identify anomalies indicated by the ground-based sensors is due to variations in anomaly density across the 10 different areas at which 100 percent coverage is available for both sensor modes.

The geostatistical estimation algorithm known as kriging is used to estimate geophysical anomaly density at unsampled locations using the limited coverage transect data collected with the ground-based sensors. These estimates are compared to the results of the 100 percent coverage provided by the helicopter-mounted sensors. Comparisons are made using the distribution of estimated and measured anomaly density values at the scale of 25x25m cells within each of the 10 areas. Additionally, comparisons are made between the areas estimated as having high anomaly density, as would be considered a target area in a site characterization, and those areas classified as high anomaly density based on the 100 percent helicopter-based survey coverage. These comparisons showed the kriging estimates successfully identifying the areas of high anomaly density observed in the measured data. The measured anomaly densities from the 100 percent coverage helicopter surveys are 13 and 14 percent of the anomaly density estimates made using the

sparse and full transect data sets, respectively. This result is consistent with the range of underestimation (8 to 29 percent) of the total anomalies by the helicopter-based sensors relative to the ground-based sensors in areas where both have 100 percent coverage. The shape of estimated anomaly density distributions shows that the kriging estimates better represent the lower density regions than do the helicopter based data, which underestimate densities in these areas. For higher density regions, those above 200 – 240 anomalies per acre, the kriged estimates and the helicopter-based measurements are more consistent.

The same kriging estimates made from the sparse and dense transect data sets are compared to the results of the 100 percent ground-survey coverage in the 10 areas where that level of coverage is available. Comparison of the distribution of the estimated and measured anomaly densities is made along with comparison of the delineation of high anomaly density areas. Results show that the kriging estimator is capable of reproducing the density anomaly distributions within each of the 10 areas, with the exception that the highest density values in each are consistently underestimated by the kriging. This result is expected as the kriging estimates are based on the ground-sensor surveys covering only 0.64 and 1.64 percent of the site area, and these limited survey areas are not expected to intercept the highest anomaly density regions of each 100 percent coverage area. The kriging estimates also provide accurate delineation of the high anomaly density areas. This result is quantified using ROC curve analysis and remains consistent across three different definitions of "high" anomaly density: 10, 20 and 30 anomalies per acre. Results of this comparison show that both the location of a high anomaly density area and the distribution of densities within that area can be accurately estimated from very limited sampling coverage using kriging. These results are especially noteworthy given the nearly two order of magnitude variation in anomaly densities across the 10 regions of 100 percent ground survey coverage that are available.

The 100 percent survey coverage by the helicopter-mounted sensors allows different simulated transect sampling patterns to be chosen and extracted from the 100 percent helicopter coverage. Six different transect patterns: 3 with north-south orientation, but different origins, as well as east-west, northeast-southwest and northwest-southeast orientations are selected. The transect spacing is kept at the original spacing distance of 156 meters. As in the case of the ground-based transect data, the kriging estimator uses the selected helicopter transects to estimate anomaly densities at all unsampled locations. These estimates are then compared to the results of the 100 percent survey and the ability of the estimates to locate areas of high anomaly density is assessed. High density is defined here as 12 anomalies per acre which is the ground-based density threshold value used in the WAA (60 anomalies per acre) scaled by the average reduction factor, 0.20, from ground-based to helicopter-based anomaly counts. All selected transect designs provide data sets that allow for accurate delineation of high anomaly density areas. ROC curve analysis quantifies this accuracy with areas under the ROC curves ranging from 0.86 to 0.92. Statistical testing shows these curve areas are not different at a 0.05 significance level.

In summary, the analyses described in this report demonstrate the following: 1) The helicopter mounted sensors record only 8 to 29 percent of the anomalies recorded by the ground-based sensors over the same areas; 2) Spatially varying estimates of anomaly densities derived from limited transect sampling with ground-based sensors differ from the number of anomalies recorded by helicopter mounted sensors by an expected amount. The helicopter-based anomaly densities are approximately 14 percent of the values estimated using the ground-based sensors as inputs; 3) Geostatistical estimation using ground-based sampling transects covering only 0.64 and 1.64 percent of the site area is capable of accurately estimating both the distribution of anomaly densities and the spatially varying patterns of high and low anomaly density in 10 regions containing 100 percent survey coverage; and 4) The accuracy of geostatistical estimation in delineation of high anomaly density areas across the site is insensitive to the location and orientation of the survey transects used as input to the estimation.

This work verifies the approach of using limited transect sampling, less than 2 percent of the site area, with geostatistical estimation to accurately estimate the number of geophysical anomalies and their spatial distribution across a wide-area site. A practical implication of this result is that at sites with conditions similar to those at PPBR where helicopter survey rates of 400 acres per day can be achieved, an accurate estimate of the geophysical anomaly density distribution for an entire 40-50,000 acre area could be achieved in just one or two days. The kriging process is fast enough that when new data become available, the overall anomaly density map could be updated in an hour or two and then used by the site characterization team to focus the next phase of sampling.

1. Introduction

The efficient characterization and remediation of sites that are potentially contaminated with unexploded ordnance (UXO) remains a high priority for the Department of Defense (DoD). Recent estimates of the amount of land that is potentially contaminated with UXO are as high as 10 million acres (> 4 million hectares). This total land area is comprised of as many as several thousand individual sites. Characterization efforts to date have shown that at a typical site, the UXO contamination is concentrated in a small portion of the site, often only 10-20 percent of the entire site area. Therefore, efficient site characterization is focused on identifying the location and extent of these smaller areas within the larger area of the site. Toward this goal, Pacific Northwest National Laboratory (PNNL) and Sandia National Laboratories (SNL) have developed efficient and defensible, statistically-based approaches for UXO site characterization.

PNNL and SNL have developed statistical algorithms to create transect designs based on desired Data Quality Objectives (DQO) and then identify potential target areas based on the survey transects. The transect design tools provide a statistically defensible method to use transect survey data that covers only a small proportion of the total study area (i.e. 1% to 3%) to identify target areas of a desired size, shape and anomaly density. Target area density estimates, probability estimates, and density flagging routines are applied after the data are gathered from the established transect design to separate potential target areas from areas that require no further remediation.

The ESTCP program established several demonstrations of UXO site characterization technologies under a Wide Area Assessment (WAA) project. This report covers an investigation into the performance of these statistically-based tools to the Pueblo Precision Bombing and Pattern Gunnery Range (PPBR) located in South-East, Colorado. Specifically, this report compares geostatistical based estimates of anomaly densities developed from statistically based transect surveys of limited spatial coverage, to anomaly density values derived from air and land based full-coverage magnetometer surveys. These comparisons provide useful information regarding the effectiveness of using survey transect data in site characterization.

This report is organized as follows: Section 1 discusses introductory and background information; Section 2 presents the PPBR WAA site; Section 3 discusses the magnetometer systems used to collect magnetic anomaly data and the data sets used in this study; Section 4 presents the comparisons between the geostatistical estimates of magnetic anomaly density and the full-coverage, field measured values; Section 5 presents an analysis of transect design sensitivity; and Section 6 provides a summary of the study conclusions.

2. Pueblo WAA Site Information

The WAA study site used in this study is a small part of the larger Pueblo Precision Bombing and Pattern Gunnery Range #2 as shown in Figure 1. It covers approximately 7300 acres and is located about 20 miles south of La Junta, Colorado, in Otero County. The range was issued a Certificate of Clearance (surface only) in 1946 for a land use of cattle grazing. The WAA study area includes two known targets, Bomb Target 3 and Bomb Target 4 located in the northern and southern portions of the WAA study site respectively (Figure 1).

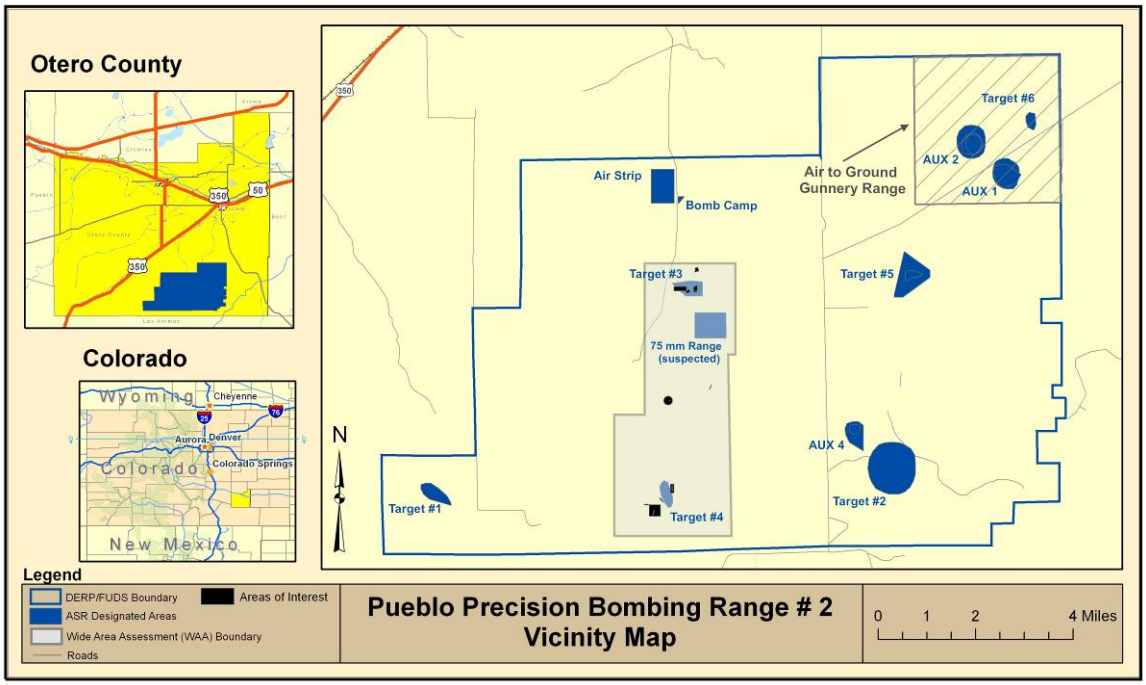


Figure 1. Vicinity map of the WAA Pueblo study area with identified features of interest.

3. Magnetometer Data

Magnetometer data collection for the Pueblo WAA consisted of both airborne and ground-based collection systems. The airborne system was helicopter based and was used to collect full-coverage data across approximately 75% of the study area. The ground-based system consisted of a towed magnetometer array. This system was used to collect full-coverage data over small areas, and, as part of a separate survey, to collect evenly spaced, statistically-planned transects across a large area of the site. Because of limiting topography and vegetation, neither magnetometer system was able to fulfill 100% of their planned data collections (Versar, 2007).

The helicopter magnetometry system consisted of a series of magnetometers mounted in a cross-track configuration forward of the aircraft body (Figure 2). Flight height during data collection is approximately 2-to-3 meters (Figure 3). The aircraft was flown in a series of parallel adjacent swaths in order to obtain a full-coverage data set. Each transect was 10.5 meters wide. Figure 4 shows the helicopter magnetometry collection area for the Pueblo WAA study area and the detected magnetic anomalies.

Figure 2. Helicopter-based magnetometer system used at the Pueblo WAA.

Figure 3. Example of helicopter magnetometry collection at the Pueblo WAA.

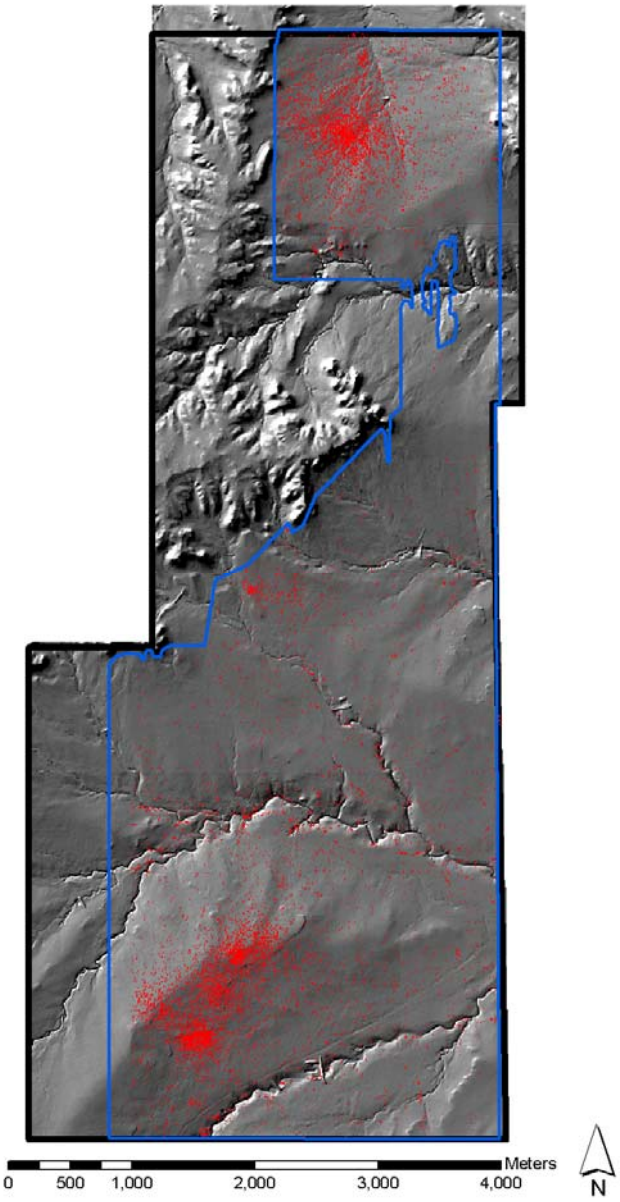


Figure 4. Helicopter magnetometry coverage for the Pueblo WAA. Blue outline indicates area for which helicopter magnetometry data was available. Black outline shows the boundary of the Pueblo WAA study area. Red dots indicate individual magnetic anomalies detected in the helicopter magnetometry data.

The ground-based magnetometer data were collected using a towed-array system which was pulled behind an all terrain vehicle (Harbaugh et al., 2006). Figure 5 shows this system during data collection at the Pueblo site. This system uses a linear array of eight magnetometers spaced at 25 centimeters in the cross-track direction. This configuration gives a nominal transect width of 2 meters. This system was used to collect individual

transects based on transect designs planned to have a high probability of intersecting target areas of a specific size. The complete magnetometer coverage collected as part of the statistically designed transect survey is shown in Figure 6. The transect coverage shown in Figure 6 represents the summation of all the transects collected using the PNNL/SNL statistically-based target identification procedures. During application of these evaluation procedures, individual transect subsets were extracted from the total transect coverage and used to investigate different transect scenarios. Analyses in this report will focus on two of these scenarios. These two scenarios represent the lowest and highest transect density subsets used during the statistical target identification analysis and will be referred to as the "sparse" and "dense" transect scenarios respectively. The transect coverage represented by these two scenarios is presented in Figure 7 (ESTCP, 2006a).

Figure 5. Ground-based magnetometer system used at the Pueblo WAA site. Magnetometer sensors are located along the structural framing mounted below the mid-section of the trailer system.

Table 1 and Table 2 present some basic characteristics of the sparse and dense transect data. Table 1 lists the total area covered by the transects and the total number of magnetic anomalies detected for both scenarios. Table 2 presents information for the two scenarios after up-scaling of the raw anomaly location data. Included in Table 2 are anomaly density values in anomalies per acre. All anomaly density values discussed in this report will be in units of anomalies per acre (ApA). The up-scaling procedure operates by computing the average anomaly density for a rectangular window as it is translated along the path of the transects. This rectangular window was oriented parallel to the cardinal directions and measured 325 m in the north-south dimension and 125 m in

the east-west dimension. Details regarding this up-scaling procedure can be found in ESTCP, 2006a. The resulting anomaly density distribution is much smoother than the original raw or unsmoothed magnetic anomaly data. The smoothed, up-scaled data will be used for all subsequent geostatistical analyses.

In addition to the individual transect surveys, the ground-based magnetometer system was also used to collect full-coverage data at selected site locations. The ground-based full-coverage surveys were performed to better define the anomaly density distribution around known target areas and to characterize the background anomaly density. The locations and naming conventions of these ground-based full-coverage areas are shown in Figure 8. Basic characteristics of the magnetic anomaly data collected from these areas are listed in Table 3. Further information regarding collection of the full-coverage areas using the ground-based magnetometer system can be found in Harbaugh et al., 2006.

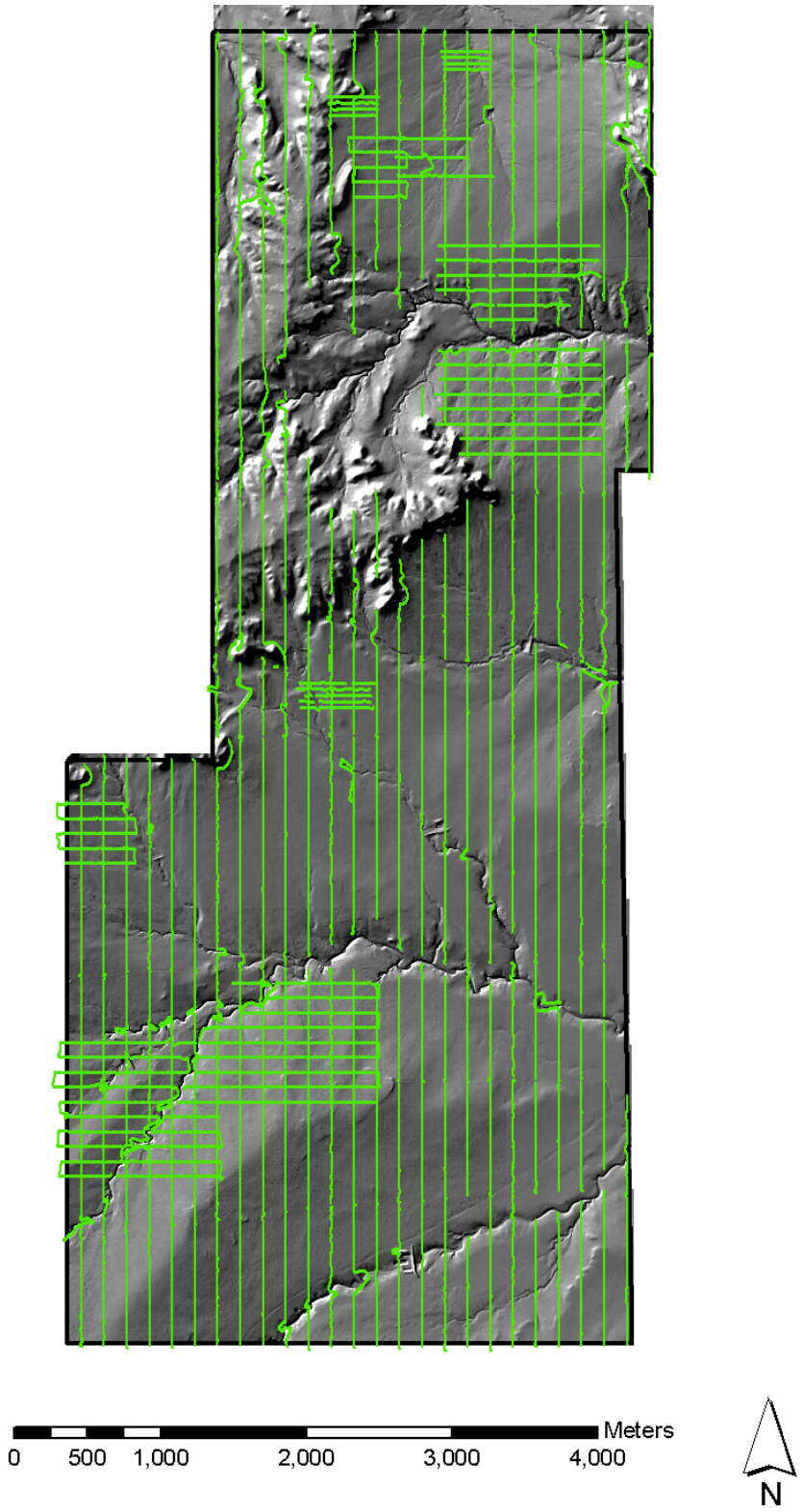


Figure 6. Ground-based magnetometer transect coverage (green lines) used for statistically-based target delineation and anomaly density estimation for the Pueblo WAA site.

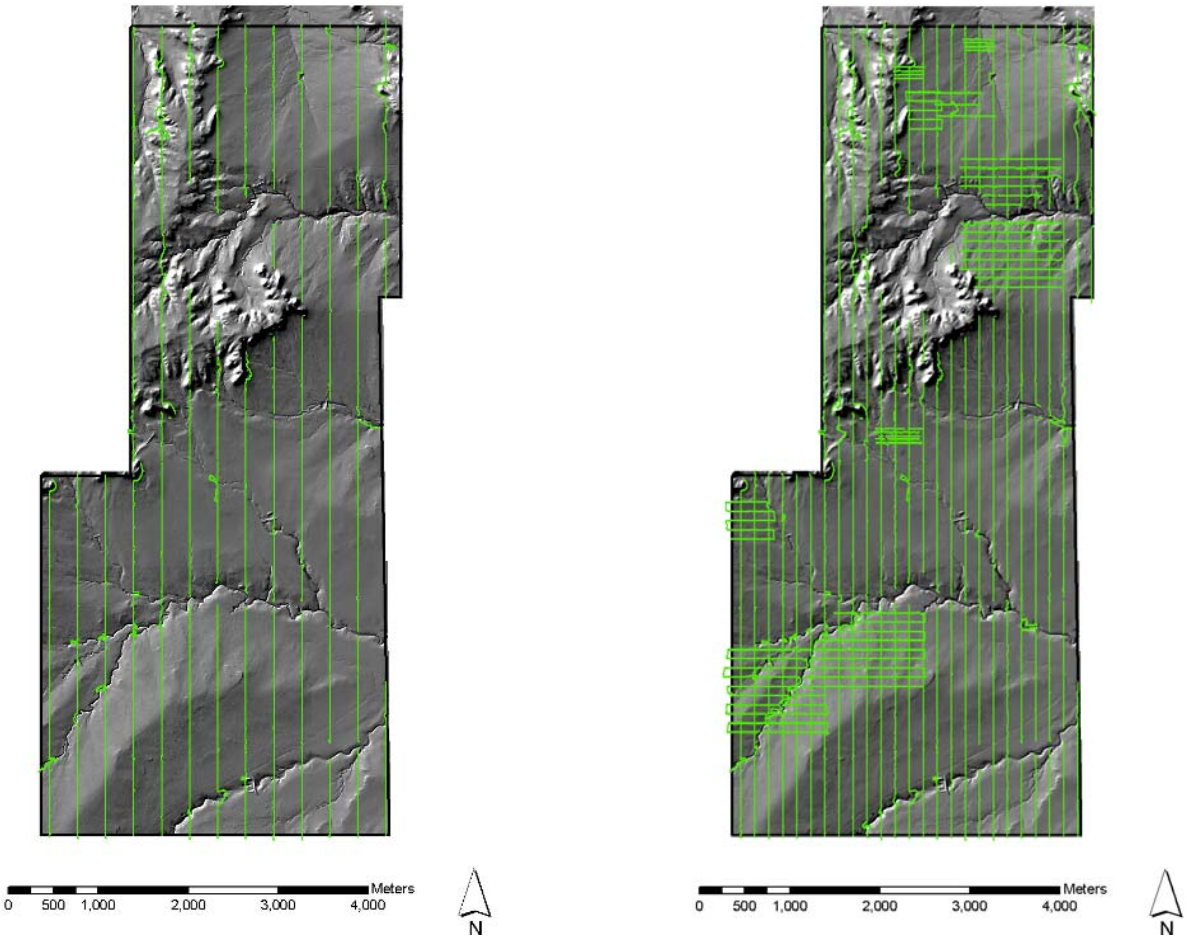


Figure 7. Sparse (left plot) and dense (right plot) ground-based transect coverage. These represent the lowest and highest (respectively) transect density scenarios used in the original statistical target identification report.

Table 1. Coverage statistics and anomaly counts for sparse and dense transect scenarios for the raw transect anomaly data. Percent coverage represents the portion of the entire study area covered by the ground-based transects.

	Sparse Transect Scenario	Dense Transect Scenario
Coverage (m^2)	192,151	491,245
Percent Coverage	0.64%	1.64%
Anomaly Count	672	1881
Nominal Transect Spacing	312 m	156 m north-south 102 m east-west

Table 2. Basic statistics for anomaly density (anomalies per acre) for the sparse and dense transect scenarios for the up-scaled data used in the statistical target analysis.

Data Set Value	Sparse Transect Scenario	Dense Transect Scenario
n (anomaly count)	3,966	9,600
Maximum	280.4	311.2
Median	6.2	6.2
Minimum	0.0	0.0
Mean	13.9	15.5
Standard Deviation	31.5	31.8

Table 3. Basic characteristics of the full-coverage areas surveyed with the ground-based magnetometer system. Area is in acres, anomaly density in ApA.

Area ID	Area	Anomaly Count	Anomaly Density
1A	33.6	169	5.0
1B	33.9	245	7.2
1C	38.8	1095	28.2
BT4	11.0	938	85.3
2A	31.0	148	4.8
2B	36.8	83	2.3
3A	35.4	2112	59.7
3B	36.3	520	14.3
3C	35.7	207	5.8
Simmons	85.0	72	0.8

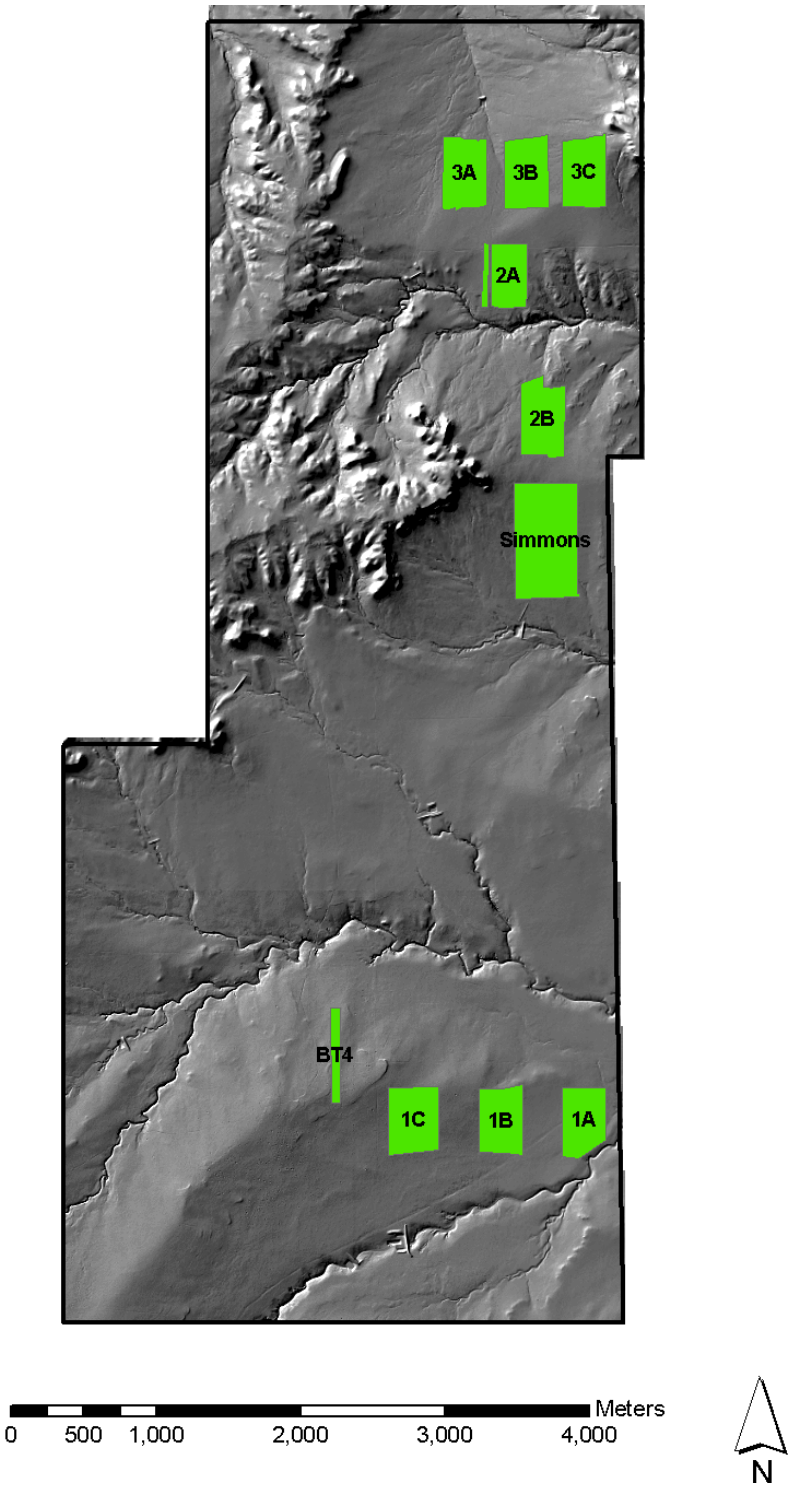


Figure 8. Ground-based magnetometer full-coverage areas (green polygons) for the Pueblo WAA site. Text labels identify individual full-coverage areas.

4. Comparison of Transect-based Kriging Estimates and Full-coverage Data Sets

This section will compare the geostatistical based estimates of anomaly densities from the statistically based transect surveys to those derived from full-coverage magnetometer surveys. The geostatistical estimates of anomaly densities derived from the limited transect survey data (<2% of site sampled) collected using the ground-based magnetometer system provided density values for the entire PPBR study area. The ability to develop anomaly density estimates for the entire site, solely from limited transect data, provides for substantial savings in time and cost, and the full-coverage magnetometer data sets provide a means to evaluate how effective this technique is. The geostatistical anomaly density estimates will be compared to both ground-based and helicopter-based full-coverage data sets.

4.1. Kriging Density Estimates

Geostatistical (kriging) estimates of magnetic anomaly density have previously been developed from the ground-based transect data for the entire PPBR site (ESTCP, 2006a). These estimates were developed using kriging to extrapolate information away from the transect locations to unsurveyed locations of the site domain. Specifically, Ordinary Kriging (OK) (Isaaks and Srivastava, 1989) was employed using anomaly density data from the up-scaled magnetometer transect survey data. Variograms for use in the kriging were developed for each specific transect scenario. Details regarding development of the geostatistical estimates can be found in ESTCP (2006a).

A wide variety of magnetometer transect sampling scenarios were investigated during the original development of the kriging estimates (ESTCP, 2006a) for the PPBR site. In order to represent the full range of these scenarios with in-depth analysis, the comparisons presented here will concentrate on kriging estimates developed from the highest and lowest transect density scenarios (Figure 7). It is the kriging estimates of magnetic anomaly density developed using these transect data sets that will be compared against the full-coverage data sets.

4.2. Helicopter-based Full-coverage Data Set Comparison

In this section the kriging estimates of anomaly density developed from ground-based magnetometer transect data will be compared to the helicopter-based full-coverage data set. This section also includes discussion of the pre-processing of the helicopter magnetometer data that was necessary to prepare it for use in the comparison, and a discussion on the challenges of comparing helicopter and ground-based magnetometer data sets.

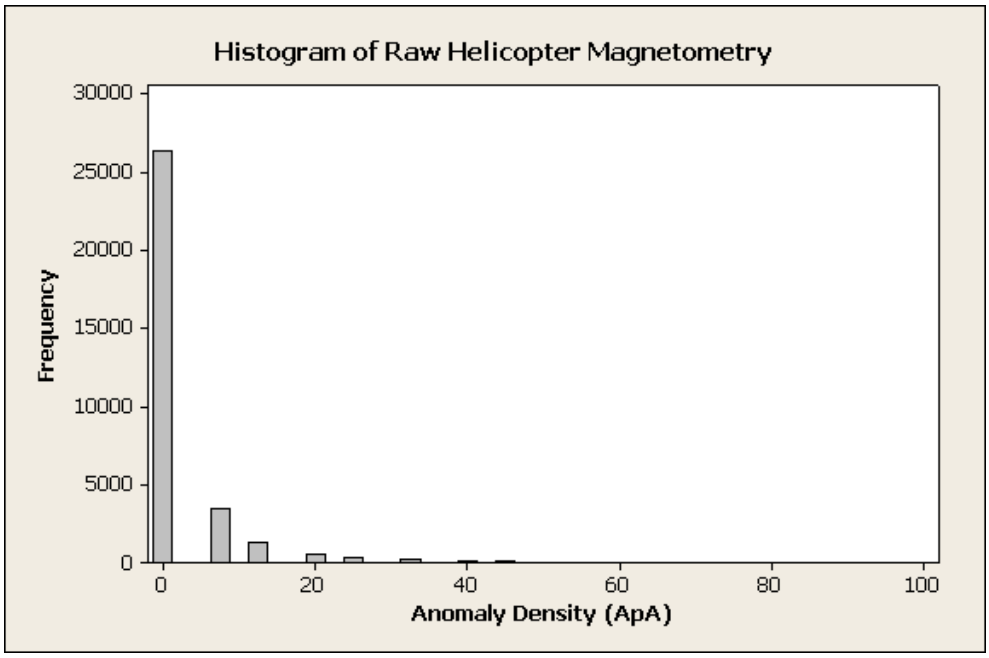
4.2.1. Helicopter data set preparation

In order to compare the full-coverage helicopter magnetometer data to the kriging estimates of anomaly density it was necessary to transform the anomaly locations identified in the helicopter magnetometry to area-based density values. This was done by

overlaying a 25 x 25 m matrix of grid cells over the helicopter magnetometry anomaly data set. The grid used for this process was located and sized so as to overlay the grid of kriging estimates in a one-for-one basis. Having a one-to-one relationship between the helicopter magnetometer data and the kriging estimates allows for direct spatial comparison between the two data sets. Because the overall coverage for the helicopter magnetometry is smaller than that for the ground-based transect survey, there are some areas of the kriging estimates developed from the transect survey data that will be excluded from the analysis; only areas with both helicopter magnetometry and kriging estimates developed from the ground-based transect data will be included in the comparisons.

Because of the discrete nature of the magnetic anomaly detections and the continuous nature of the density estimates, a strict grid cell by grid cell computation of anomaly density is not desirable. When computed in this manner, the resulting density values are binned by the number of anomalies in a grid cell and are not continuous. This binning results in only 19 unique density values for the entire helicopter magnetometry data set. To address this issue and provide a more continuous and representative distribution of density values, a quadratic kernel function was used to compute the density of the helicopter magnetometry anomalies. The quadratic kernel function is a window averaging technique used to compute density or intensity values for discrete point processes (Schabenberger and Gotway, 2005). In this, the density of points in a grid cell is computed as a weighted average of the point density in a window of surrounding grid cells. The weighting function is symmetrical with the weighting values decreasing exponentially with increasing distance. For the helicopter magnetometry anomalies the averaging window consisted of the 8 nearest neighbors (Moore neighborhood). Figure 9 shows histograms of the helicopter magnetometry data before and after application of the quadratic kernel. Application of the quadratic kernel results in more continuous distribution of density values and resolves the binning issue seen in the raw helicopter magnetometry data.

Table 4 provides a listing of the basic characteristics of the helicopter magnetometry density data from both the raw and quadratic kernel (smoothed) data sets. As this table shows, the raw density values are limited to only 19 unique values (including 0). This limited range in values is not representative of the expected density distribution and could lead to inaccuracies when compared to the more continuous kriging estimates. The density values computed using the quadratic kernel have a continuous distribution which provides a better data set to compare against the kriging estimates. The smoothing of the quadratic kernel has reduced the maximum density value and the standard deviation compared to the raw density data, but has had only minor impact on the mean value. Figure 10 presents a spatial comparison between the raw and smoothed anomaly density data.



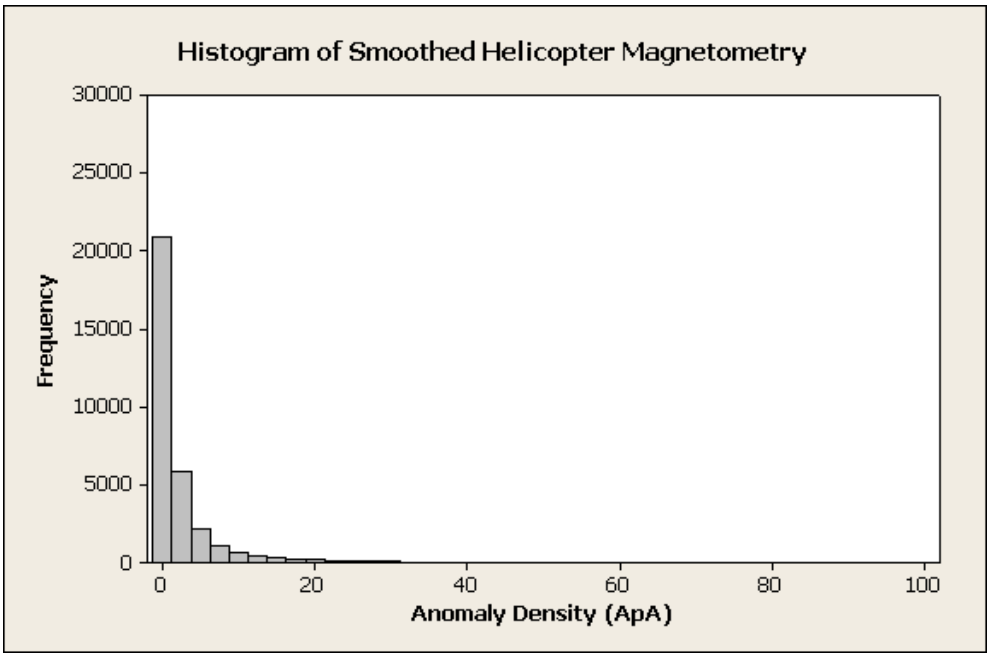


Figure 9. Histograms showing distributions of raw helicopter magnetometry anomaly densities (top plot) and smoothed helicopter magnetometry anomaly densities (bottom plot) for the PPBR site.

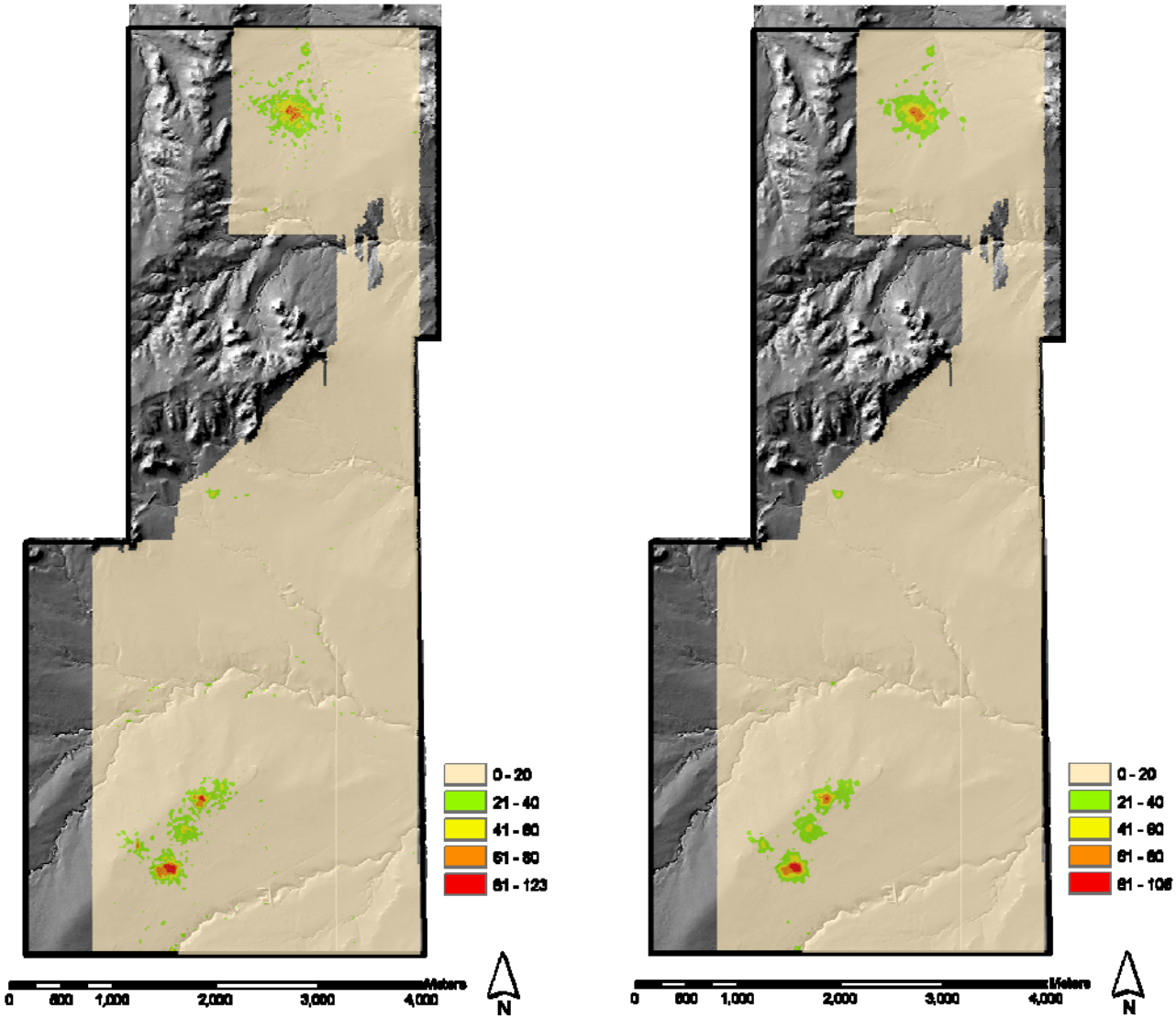


Figure 10. Comparison of raw (left figure) and quadratic kernel smoothed (right figure) helicopter magnetometry density data. Values are in ApA.

Table 4. Basic characteristics of the raw and quadratic kernel (smoothed) helicopter anomaly density data sets. The S.D. column reports standard deviation. All values are in ApA.

	Minimum	Mean	Maximum	S.D.	Unique Values
Raw Density	0.0	2.5	123.0	7.5	19
Quad. Kernel	0.0	2.4	106.4	6.4	*Continuous*

4.2.2. Helicopter and ground-based data comparison

The ground-based and helicopter-based magnetometer surveys collected at the PPBR differ significantly in their ground-to-sensor distances. This difference, along with slightly different magnetic anomaly identification and processing techniques used by the geophysical processing teams (Harbaugh et al., 2006, ESTCP, 2006b), means that the number of anomalies detected by the two systems over the same area will differ. This difference is most clearly seen in a comparison between the full-coverage areas collected with the ground-based system and the helicopter magnetometry data. Table 5 lists the anomaly counts for each system for each of the full-coverage areas covered by the ground-based system (see Figure 8). This table was developed by tabulating the number of helicopter magnetometry anomalies (raw anomaly locations) falling within each of the full-coverage areas covered by the ground-based system. Table 5 clearly shows the lower number of anomalies detected by the helicopter-based system for the same coverage areas. The ratio of helicopter-to-ground based anomaly detections range from 0.29 to 0.08, with an area-averaged mean of 0.2. This indicates that for these areas, the helicopter-based system data set contains on average about 1/5 of the anomalies contained in the ground-based system data set.

Table 5. Summary of anomaly counts and densities for ground-based (GB) and helicopter-based (HB) magnetometers for ground-based full-coverage areas. HB/GB Ratio column gives ratio of helicopter-based to ground-based anomaly counts.

Area ID	Area (acres)	Ground-Based Anomaly Count	GB Anomaly Density (ApA)	Helicopter-Based Anomaly Count	HB Anomaly Density (ApA)	HB/GB Ratio
1A	33.6	169	5.1	46	1.4	0.27
1B	33.9	245	7.2	51	1.5	0.21
1C	38.8	1095	28.2	104	2.7	0.10
BT4	11.0	938	85.3	123	11.2	0.13
2A	31.0	148	4.8	31	1.0	0.21
2B	36.8	83	2.3	7	0.2	0.08
3A	35.4	2112	59.7	609	17.2	0.29
3B	36.3	520	14.3	101	2.8	0.19
3C	35.7	207	5.8	32	0.9	0.15
Simmons	85.0	72	0.8	18	0.2	0.25

4.2.3. Helicopter and kriging estimate comparison

Because of the differences in detection rates between the helicopter and ground-based magnetometer data sets (Table 5), a direct comparison between the helicopter-based anomaly densities and the kriging estimates based on ground-based magnetometer transects is problematic. Table 6 presents basic statistics for anomaly density from the helicopter magnetometer data and the two kriging estimates. Similar to Table 5, the estimates from the ground-based transect data have higher density values than that computed for the helicopter-based data. This again reflects the differences in detection rates between the helicopter and ground-based magnetometer systems. The ratio of total detected anomalies between the helicopter and the sparse and dense transect kriging estimates (Table 6) are 0.14 and 0.13 respectively. These values fall within the range of those observed when comparing the full-coverage helicopter and full-coverage ground-based data sets (Table 5). This analysis provides a basic check that the kriging density estimates are comparable to the full-coverage ground-based magnetometer surveys.

Table 6. Comparison of basic magnetic anomaly density statistics for the helicopter magnetometry and kriging results. HB/KE ratio lists the ratio of total anomaly counts between the helicopter magnetometry and kriging estimates. All measures are in ApA and the kriging estimated values are based on a 25 m grid-cell spacing.

	Helicopter Magnetometry	Sparse Transect Kriging	Dense Transect Kriging
Mean	2.54	17.88	18.41
Std. Dev	6.38	34.72	37.26
Minimum	0.00	0.00	0.00
1st Quartile	0.00	1.29	1.90
Median	0.10	6.88	6.94
3rd Quartile	2.41	17.38	17.24
Max	106.44	280.39	311.24
Total Anomalies	12,181	89,276	91,897
HB/KE Ratio	*NA*	0.14	0.13

Although the differences in detection rates between the helicopter magnetometry and the ground-based transect data used for the kriging estimates, limits the comparison of the raw density data a general comparison of the distribution of anomaly density values can be shown using a quantile-quantile (Q-Q) plot. A Q-Q plot shows a direct comparison between two ranked data sets. If the distributions of the two data sets are similar, the ranked data will plot along a 45-degree line originating at the origin. Deviations from this line show differences between the two distributions. Figure 11 shows Q-Q plots with ranked smoothed helicopter-based anomaly densities on the Y-axes, and the ranked sparse and dense transect kriging estimates of anomaly density on the X-axes. Each point in the Q-Q plot represents a data point from the spatial grid of density values. The location of each point is determined by the pair of anomaly density values (helicopter and kriging) for that particular ranking (increasing anomaly density values) in the data set.

As Figure 11 shows, the Q-Q plots for both comparisons fall significantly below the 45-degree line indicating substantial differences between the helicopter-based data and the ground-based transect kriging estimates. The differences between the helicopter and kriging density curves are likely due to the smoothing affect of the kriging process, and the relative higher sensitivity of the ground-based magnetometer system. The higher sensitivity of the ground based magnetometer results in a smaller percentage of transect area with very low or zero anomaly densities compared to helicopter magnetometry collected over the same area. This lower percentage of small density values are then integrated into the kriging results which are then reflected in the Q-Q plots shown in Figure 11. Conversely, the relative lower sensitivity of the helicopter-based system, due to increased ground-to-sensor distance, results in a larger number of lower anomaly density values. This is reflected in the Q-Q plot falling below the 45-degree line when the data are plotted as in Figure 11.

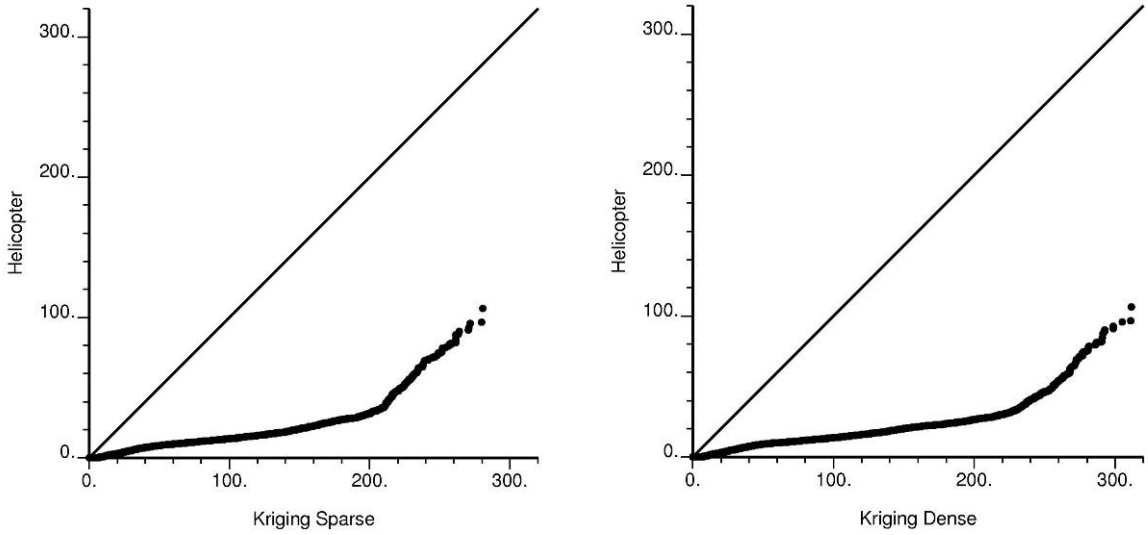


Figure 11. Quantile-quantile plots comparing the distribution of anomaly density values from helicopter magnetometry, and sparse (left) and dense (right) ground-based transect kriging estimates. All values are in ApA.

For both the sparse and dense transect kriging results, the Q-Q plot (Figure 11) shows a unique bend around the 200-240 ApA level of the kriging results axis. Initially (from 0 to 200-240 ApA), the Q-Q points define a line with a slope significantly different from the 45-degree line. This indicates that the differences between the two distributions are increasing with increasing anomaly density. Then, at about 200-240 ApA on the kriging axis (~40 ApA on the helicopter axis), the Q-Q points make an abrupt turn and plot parallel to the 45-degree line. This indicates that the two distributions (helicopter and kriging anomaly densities) now follow a similar distribution shape and that they differ by only an offset in location (magnitude). The Q-Q plots indicate that the shape of the helicopter magnetometry and kriging estimate density distributions are markedly different at lower density values, but are similar at higher values with only a shift in magnitude. Since this bend in the Q-Q plot occurs at high anomaly densities which would likely be associated with former target locations, one possible explanation for the

bend in the Q-Q plot would be a general increase in the mean-size of the metal fragments generating the magnetic anomalies. An increase in the mean fragment size may result in more similar detection rates between the ground-based and helicopter magnetometry systems resulting in more similarly shaped anomaly density distributions.

A more direct comparison between the helicopter full-coverage data and the kriging estimates was obtained by applying a density threshold to the helicopter full-coverage data to simulate the delineation of High Anomaly Density Areas (HADA) based on measured anomaly densities. HADA are often indicative of former munitions range target locations, and so are important in the characterization of these sites. Locations with anomaly density at or above the threshold would be defined as HADA locations; those below the threshold would be considered as background. This same process can then also be applied to the kriging estimates. This allows a more direct comparison by transforming the data sets from a continuous measure (anomaly density) to a binary data set (HADA, non-HADA).

To define the high-density areas, a threshold value was applied to the quadratic kernel smoothed, full-coverage helicopter magnetometer data set (Figure 10). A representative value of 12 ApA was used as the threshold for defining the HADA in the full-coverage helicopter magnetometer data; any location with an anomaly density at or above this threshold was defined as an HADA for the purposes of this analysis. From a range of possible values, the 12 ApA HADA threshold value was chosen as representative based on previous values chosen for the delineation threshold at the PPBR (ESTCP, 2006a) during the WAA. For the WAA, a delineation threshold of 60 ApA was chosen based on results from ground-based magnetometer transect data. Scaling this value (60 ApA) by 0.2 (see Table 5) to account for the reduced sensitivity of the helicopter-based system, gives the representative threshold value of 12 ApA. The 12 ApA HADA threshold was chosen for use only in this analysis, and should not be interpreted as a final delineation threshold value for the PPBR.

To investigate how well the anomaly density values developed from the transect-based kriging estimates match the HADA identifications made from the helicopter full-coverage data, Receiver Operating Characteristic (ROC) curves were developed for the sparse and dense transect kriging estimates. ROC curves are plots showing the relationship between true positive and false positive ratios for various threshold levels. For this analysis, true values are defined as those grid cells in the helicopter full-coverage data with an anomaly density at or above the HADA threshold value of 12 ApA. These will be the basis to compare the results from the transect kriging estimates to determine the false positive and true positive ratios.

The true positive ratio for the ROC curves was determined by identifying HADA in the kriging estimates as all grid cells at or above a given testing threshold. Then each grid cell in the kriging estimate was compared to the corresponding cell in the helicopter full-coverage data; if both were positive (kriging estimate above testing threshold and helicopter data above the HADA threshold) that cell would represent a true positive; if the estimate showed positive but the helicopter full-coverage data showed negative

(density value below HADA threshold) then that cell would represent a false positive. For each comparison, the number of true positives was summed and this sum divided by the number of positive cells in the helicopter full-coverage data set to obtain the true positive ratio. Similarly, the number of false positives was summed and divided by the number of false cells in the helicopter full-coverage data set to compute the false positive ratio. A series of these ratios was computed for each comparison data set by varying the testing threshold applied to the kriging estimate. As the testing threshold was increased from zero, the false positive and true positive ratios decreased from one to zero producing curves as shown in Figure 12. The testing thresholds applied to the kriging estimates ranged from 0 to 250 ApA. The lines in Figure 12 represent how the true positive and false positive ratios changed as different testing threshold values were applied to the kriging estimates. This provides an indication of how effective and efficient the model (kriged estimate of anomaly density) is in identifying the high-density areas in the helicopter full-coverage data set. Because this comparison is done on a cell-by-cell basis, it provides a good indication of how well the kriging estimates represent the spatial patterns of anomaly density recorded in the helicopter full-coverage data.

In interpreting ROC curves, curves that plot in the upper left-hand portion of the ROC curve space indicate effective predictive models. This indicates that the model has good predictive capabilities while only introducing a moderate number of false positives. If, on the other hand, one were to develop a model by randomly assigning HADA locations, then the resulting ROC curve would plot along the diagonal from (0,0) to (1,1) and would indicate that the model had no real predictive power. The closer the curve plots to the upper left-hand corner, the better the estimate of target locations. A perfect predictor of target locations would plot along the minimum X axis and maximum Y axis of the graph going through point (0,1) with a false positive ratio of zero and a true positive ratio of one.

The ROC curves in Figure 12 plot in the upper left-hand portion of the ROC curve space. Both curves in Figure 12 show an initially steep vertical rise indicating a high rate of true positive identification while maintaining a low rate of false positives. This indicates that the kriging of transect data to generate estimates of anomaly density provides an effective and efficient methodology to identifying HADA at the PPBR.

Although direct comparison is problematic, the above analyses have shown that in general the kriging estimates of anomaly density are consistent with the results from the full-coverage helicopter magnetometry data. The ratio between the helicopter measured and kriging estimated density values is within the range of values observed from field measurements at the site for helicopter and ground-based magnetometer systems. Q-Q plots comparing the kriging estimates to the helicopter data show substantial differences in the distributions of the density values, but these are likely attributed to the differences in the collection systems themselves. In the most direct comparison, the identification of HADA in both data sets, ROC curve analysis show the kriged estimates to provide a very good representation of the relative values and spatial patterns displayed in the full-coverage helicopter data.

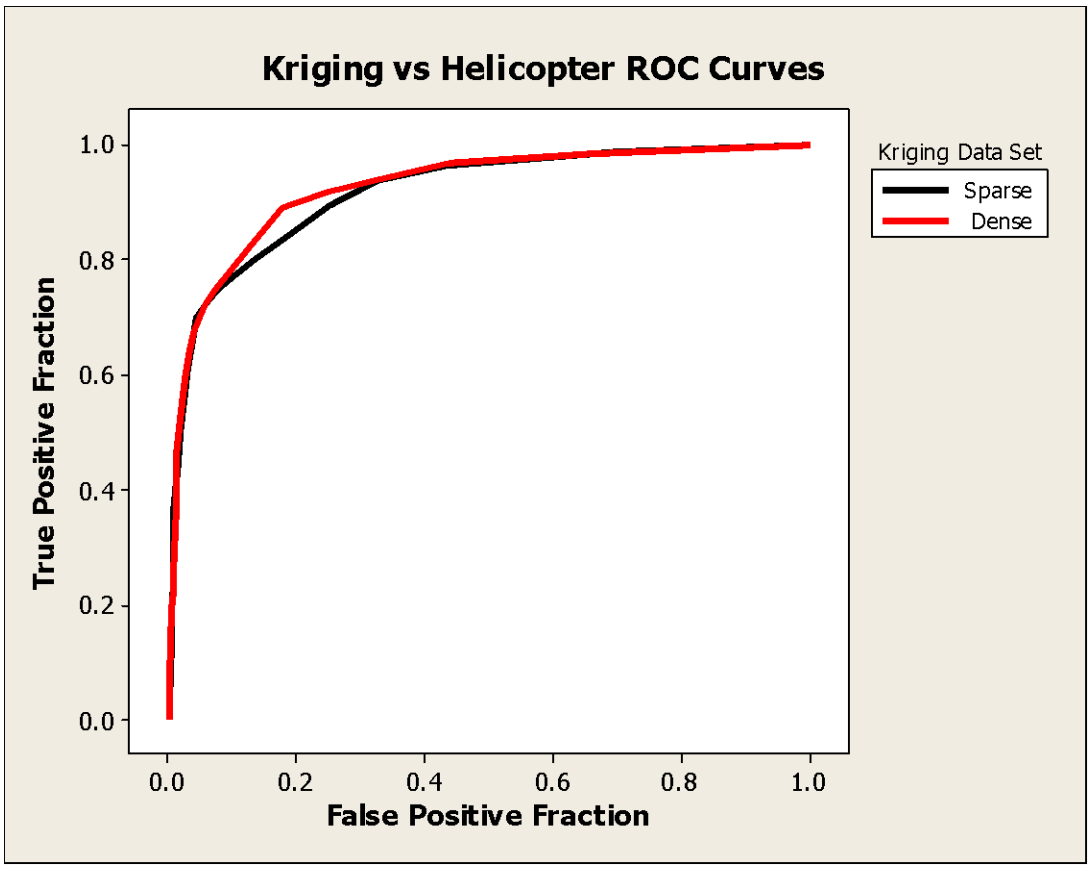


Figure 12. ROC curves comparing sparse (black line) and dense (red line) transect kriging results against the distribution of HADA in the helicopter magnetometer full-coverage data set.

4.3. Ground-based Full-coverage Data Set Comparison

This section of the report compares the kriging estimates of anomaly density to those measured at a number of small plots using ground-based full-coverage magnetometer surveys. Although of limited spatial extent, the ground-based full coverage data should provide a data set comparable to the kriging results since they were collected using the same sensor system. This should eliminate many of the issues involved with comparison of the kriging estimates to the helicopter magnetometry.

4.3.1. Data set preparation

Figure 8 shows the full-coverage areas collected using the ground-based magnetometer system. These areas are categorized into four different groups: (1) areas 1A, 1B, 1C and BT4 are associated with Target Area 4 located in the southern portion of the WAA site, (2) areas 2A and 2B are associated with the suspected 75mm area near the center of the site, (3) areas 3A, 3B, and 3C are associated with Target Area 3 at the northern end of the site, and (4) the Simmons area located near the center of the site which was chosen as representative of background anomaly density levels.

In order to compare the anomaly densities from the ground-based full-coverage areas to the kriged estimates of anomaly density developed using transect data, the anomaly data for each full coverage area was averaged using the same quadratic kernel function as applied to the helicopter magnetometry data. The grid-cell size was the same as that used for the kriging estimates (25x25 m). Table 7 shows the summary statistics compiled from all of the 10 full-coverage areas surveyed using the ground-based system compared to values compiled from the kriging estimates for the same 10 areas. Not surprisingly, the mean and maximum density values from the kriged estimates are lower than the measured values. This is due to the limited transect sampling which, for those transects within the ground-based full-coverage areas, may not necessarily intercept high anomaly density locations. Although the estimated mean and maximum density values are low, their accuracy is impressive when considering the limited area sampled, and how few anomalies were encountered in that sampling. In fact, for the sparse transect sampling scenario, 6 of the 10 full-coverage areas did not have any detected anomalies within the sample transects.

Table 7. Summary statistics for the full-coverage areas surveyed with the ground-based magnetometer system (second column), compared to those from kriging estimates using the sparse (third column) and full (fourth column) transect data. Density values use units of ApA.

	Full Coverage	Kriging (sparse)	Kriging (full)
Mean Density (ApA)	14.0	11.0	11.6
Standard Deviation	28.3	24.0	24.0
Maximum Density (ApA)	272.0	167.4	167.5
Anomaly Count	5,619	4,059	4,267
% Area Sampled	100%	0.7%	2.0%

4.3.2. Detailed comparisons

Figure 13 through Figure 16 show box and whisker plots for all of the 10 full-coverage areas comparing the ground-based full-coverage magnetometer survey results to the densities estimated via kriging of the transect sample data. These comparisons provide a direct evaluation of the transect-based kriging estimates against direct field measurements for data collected using the same geophysical survey equipment. These comparisons are important in determining how well the kriging estimates reproduce the magnetic anomaly distribution for differing density distributions and spatial locations across the WAA site.

For the following figures and discussion, density values are in anomalies per acre (ApA) as computed for each 25x25 m grid cell. In each plot, the boxes represent the interquartile (Q3-Q1) range with the median shown by a horizontal line. Vertical lines show the largest value within the upper limit (Q3+1.5(Q3-Q1)) and the smallest value within the lower limit (Q1-1.5(Q3-Q1)) with observations beyond this range shown by asterisks.

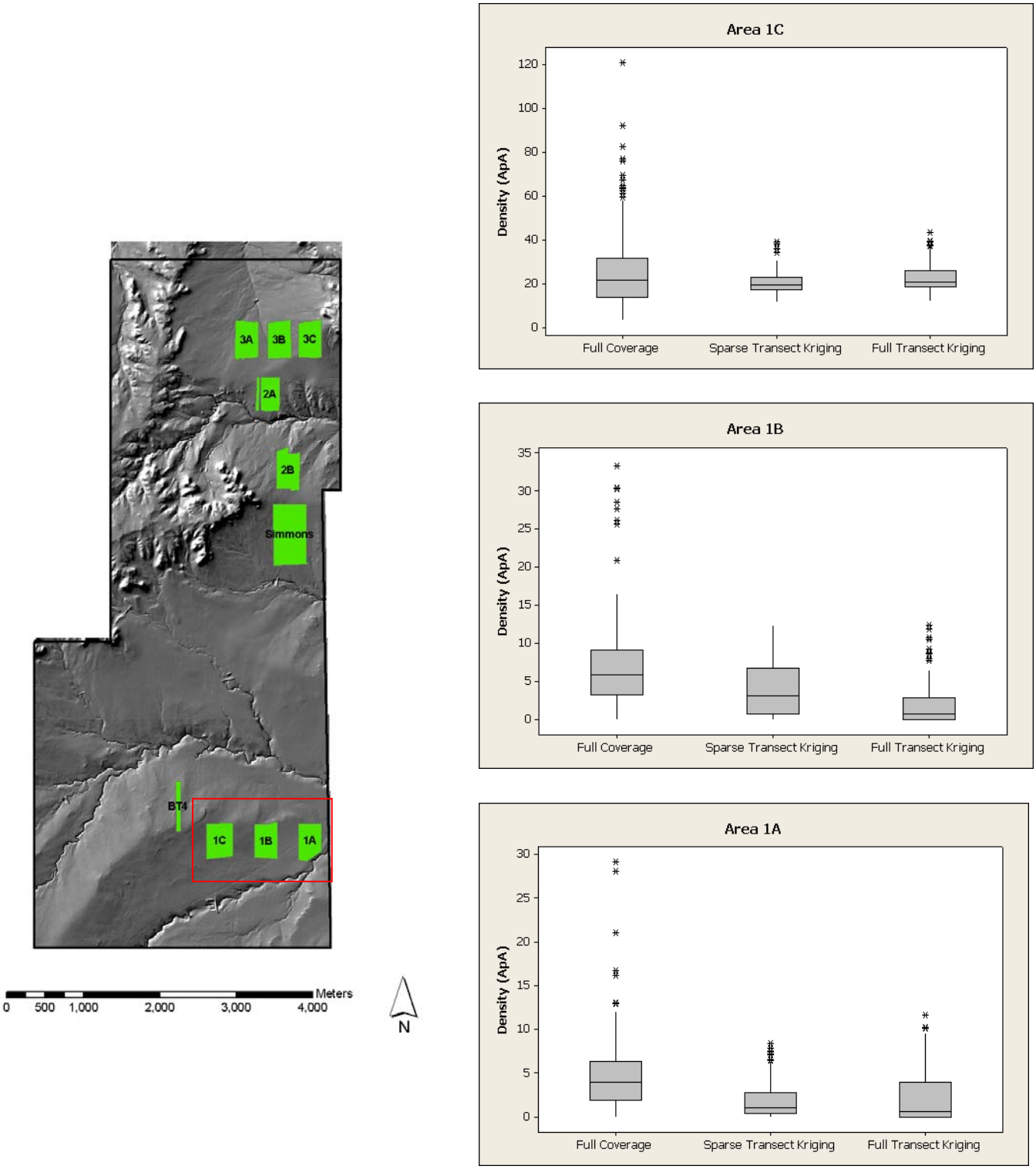


Figure 13. Box and whisker plots for full-coverage areas 1A, 1B, and 1C, comparing anomaly density values from the ground-based magnetometry survey, and the sparse and full transect kriging estimates. Figure on left shows the locations of the full-coverage areas highlighted with red rectangle.

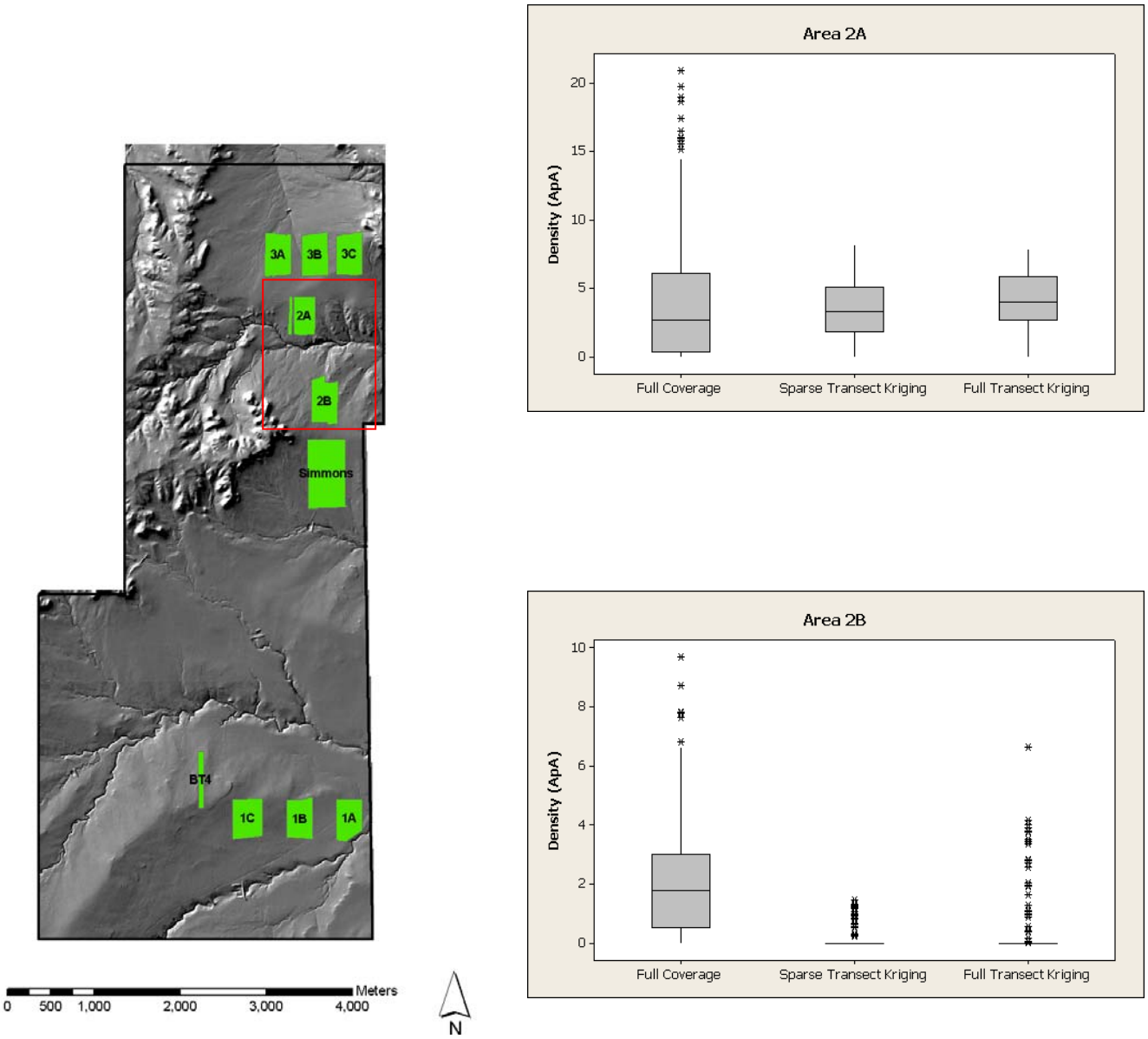


Figure 14. Box and whisker plots for full-coverage areas 2A and 2B, comparing anomaly density values from the ground-based magnetometry survey, and the sparse and full transect kriging estimates. Figure on left shows the locations of the full-coverage areas highlighted with red rectangle.

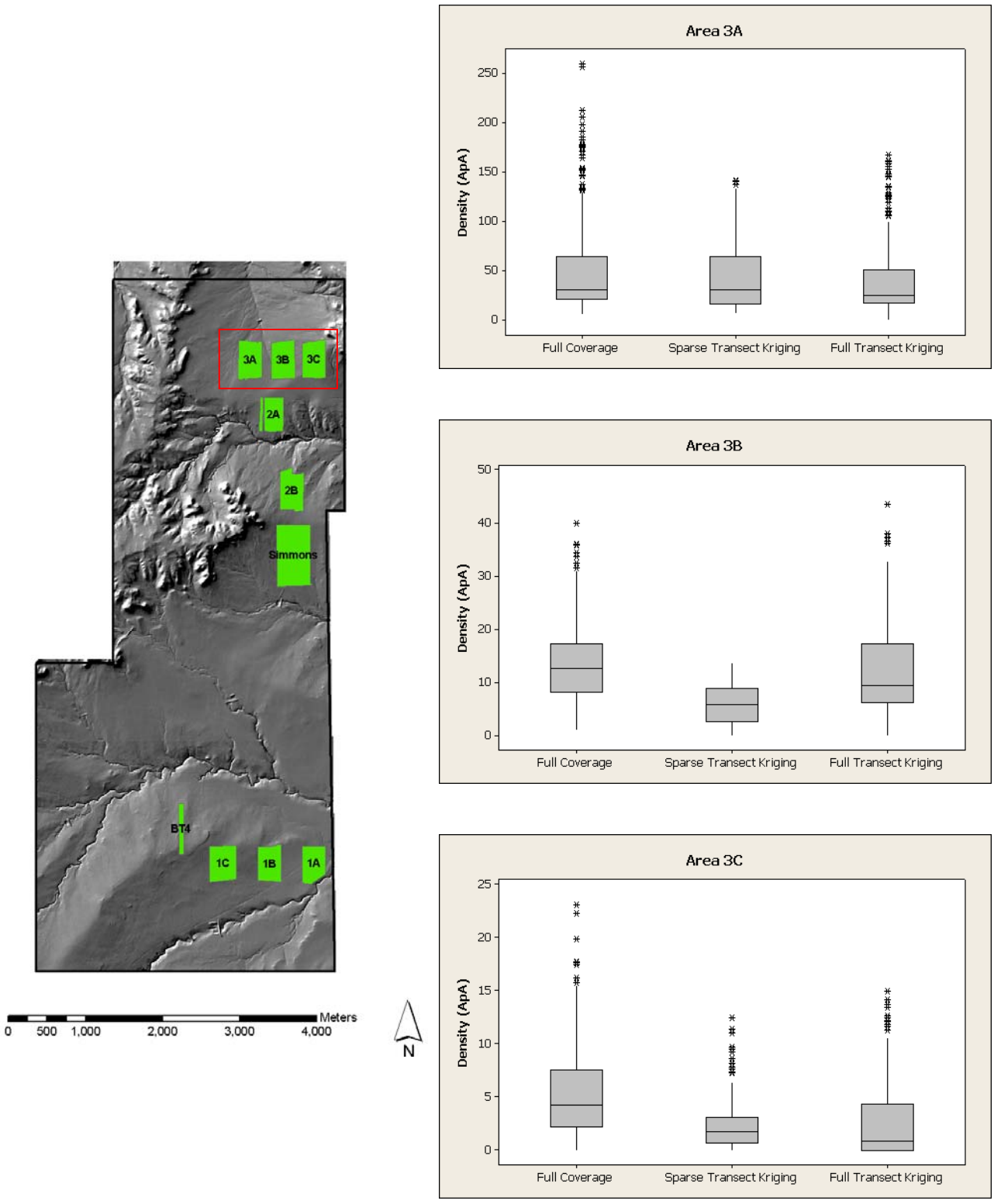


Figure 15. Box and whisker plots for full-coverage areas 3A, 3B, and 3C, comparing anomaly density values from the ground-based magnetometry survey, and the sparse and full transect kriging estimates. Figure on left shows the locations of the full-coverage areas highlighted with red rectangle.

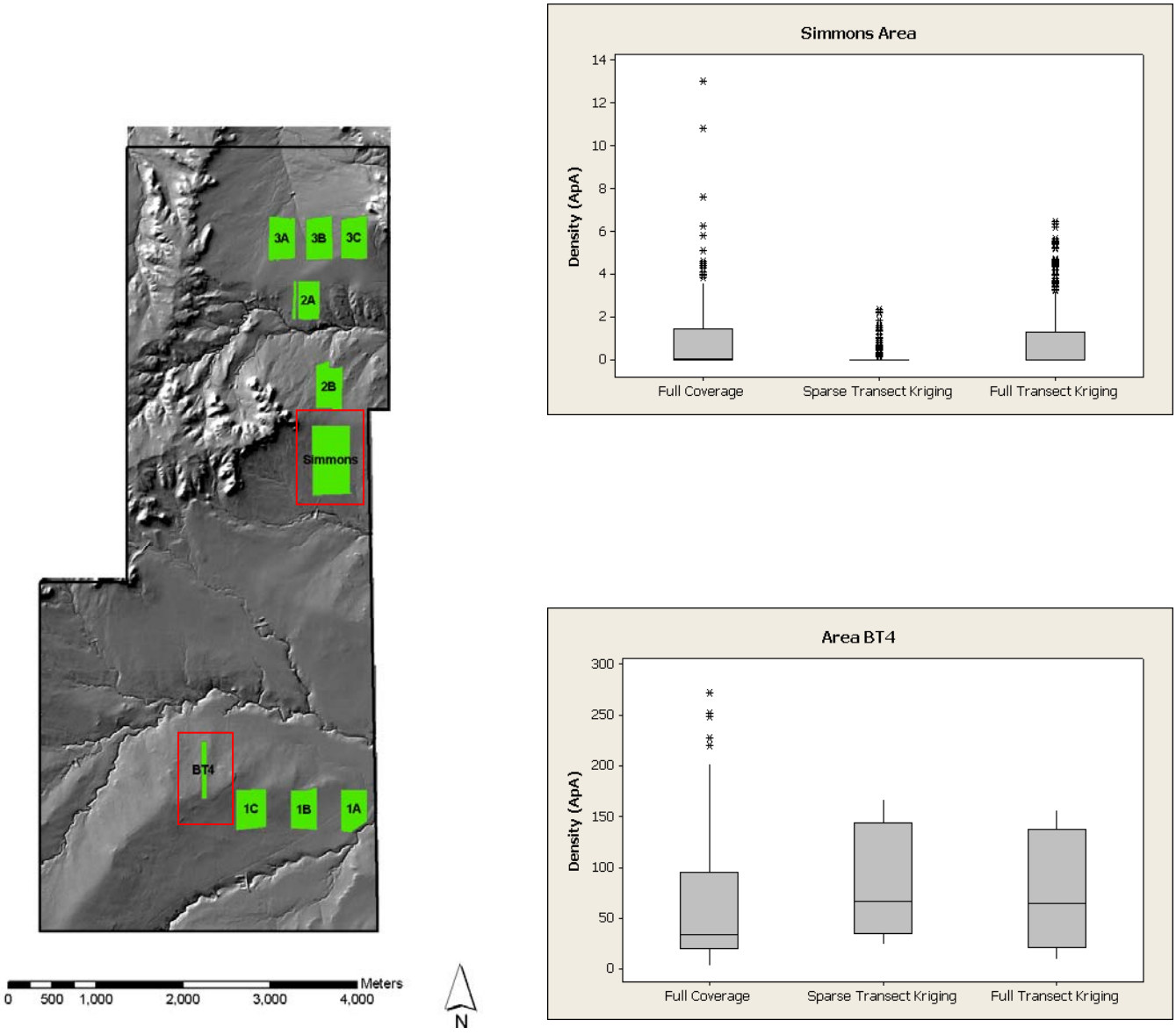


Figure 16. Box and whisker plots for full-coverage areas Simmons and BT4, comparing anomaly density values from the ground-based magnetometry survey, and the sparse and full transect kriging estimates. Figure on left shows the locations of the full-coverage areas highlighted with red rectangles.

Ground-based full-coverage areas 1A, 1B, and 1C were collected adjacent to Target 4 of the PPBR (Figure 13). The box and whisker plots in Figure 13 show how the median density for the full-coverage survey data decreases with increasing distance from the target area. This decrease is also reflected in the kriging estimates. In general, the kriged estimates show good agreement with the survey data, but with reduced upper and lower limit ranges as is expected given the smoothing nature of the kriging estimator. For the 3 full-coverage areas of Area 1, all kriging estimates have median values within 5 ApA of the measured value. Area 1C, with the highest anomaly concentrations, has the closest agreement. The median for the full transect kriging estimates are within 0.9 ApA of the full-coverage value.

Ground-based full-coverage areas 2A and 2B were located in suspected target areas. Magnetometer surveys and subsequent investigation revealed that this area was not a former target area. The low anomaly densities for this area are shown in the ground-based full-coverage density data displayed in Figure 14. The median values from the kriging estimates for these areas are in close agreement with the measured values, falling within 2 ApA for each of the areas. This indicates successful application of the anomaly density estimation procedures even in low concentration areas, which helps to limit the number of false positive target area detections.

Ground-based full-coverage areas 3A, 3B, and 3C were collected near Target 3 of the PPBR. As shown in Figure 15, the measured anomaly density decreases from area 3A to 3B to 3C which are at increasingly greater distances from Target 3. This drop in anomaly density is also reflected in the kriging results, indicating that the kriging estimate is successful in characterizing anomaly density changes near target areas. Even for the highest density area (area 3A), the median values of the kriging estimates for each area are within about 6 ApA of the measured values.

Figure 16 shows box and whisker plots for the Simmons and BT4 ground-based full-coverage areas. The BT4 ground-based full-coverage area was collected near the center of Target Area 4 and represents an area with a high magnetic anomaly density. The Simmons ground-based full-coverage area was collected in an area believed to be free of ordnance related material, and so representative of background anomaly density levels. The ranges of anomaly densities for these two areas represent the extremes observed in the ground-based full-coverage areas. The Simmons area contains the lowest densities while the BT4 area contains the highest densities (Figure 16). Considering these two areas represent opposite ends in the range of density values from the ground-based full-coverage areas, their related kriging estimates correspond well, indicating that the kriging analyses provide a good representation of the range of anomaly densities at the PPBR study site.

A spatial comparison between the ground-based full-coverage magnetometer data and the kriging estimates are shown in Figure 17 and Figure 18. These figures compare the kriging results (left hand plot) to the ground-based full-coverage anomaly density measurements (right hand plot) for the sparse (Figure 17) and dense (Figure 18) transect kriging estimates. In the right hand plot, the full-coverage anomaly density is shown as

an overlay on top of the kriging estimate (areas of lighter shading). The importance of Figure 17 and Figure 18 is in the evaluation of the spatial patterns of anomaly density from the kriging estimates. By overlaying the small plots of land-based full-coverage magnetometer data (right hand plots) onto the kriging estimates, we can see how these data compare to the kriging estimates.

The color coding for the measured and kriged anomaly densities in Figure 17 and Figure 18 is the same. Comparison of the left and right hand plots in each figure allows a visual evaluation of the kriging estimates against the measured values. Comparing the kriged results to the ground-based full-coverage measurements reveals a few isolated high-density spots not represented in the kriging estimates. These small high-density spots were not traversed by the 2-meter wide transect survey and so were not represented in the data set used in the kriging analysis. Consequently, they are not reflected in the kriging estimates. Excluding the small isolated high-density spots, and considering the limited number of transects traversing the ground-based full-coverage sites, both kriging estimates provide a good representation of the general anomaly density pattern and accurately reflect conditions in background areas.

The above comparisons have shown that the kriging estimates of anomaly density developed from limited transect data provide an effective estimate of the actual anomaly density distribution at the site. This conclusion holds for regions within the center of target areas as well as regions representative of background anomaly densities. These comparisons are particularly impressive when considering the very limited sample size (2%) used to develop the estimates. These results confirm the effectiveness of limited transect sampling with subsequent geostatistical estimation as a means of characterizing large munitions ranges.

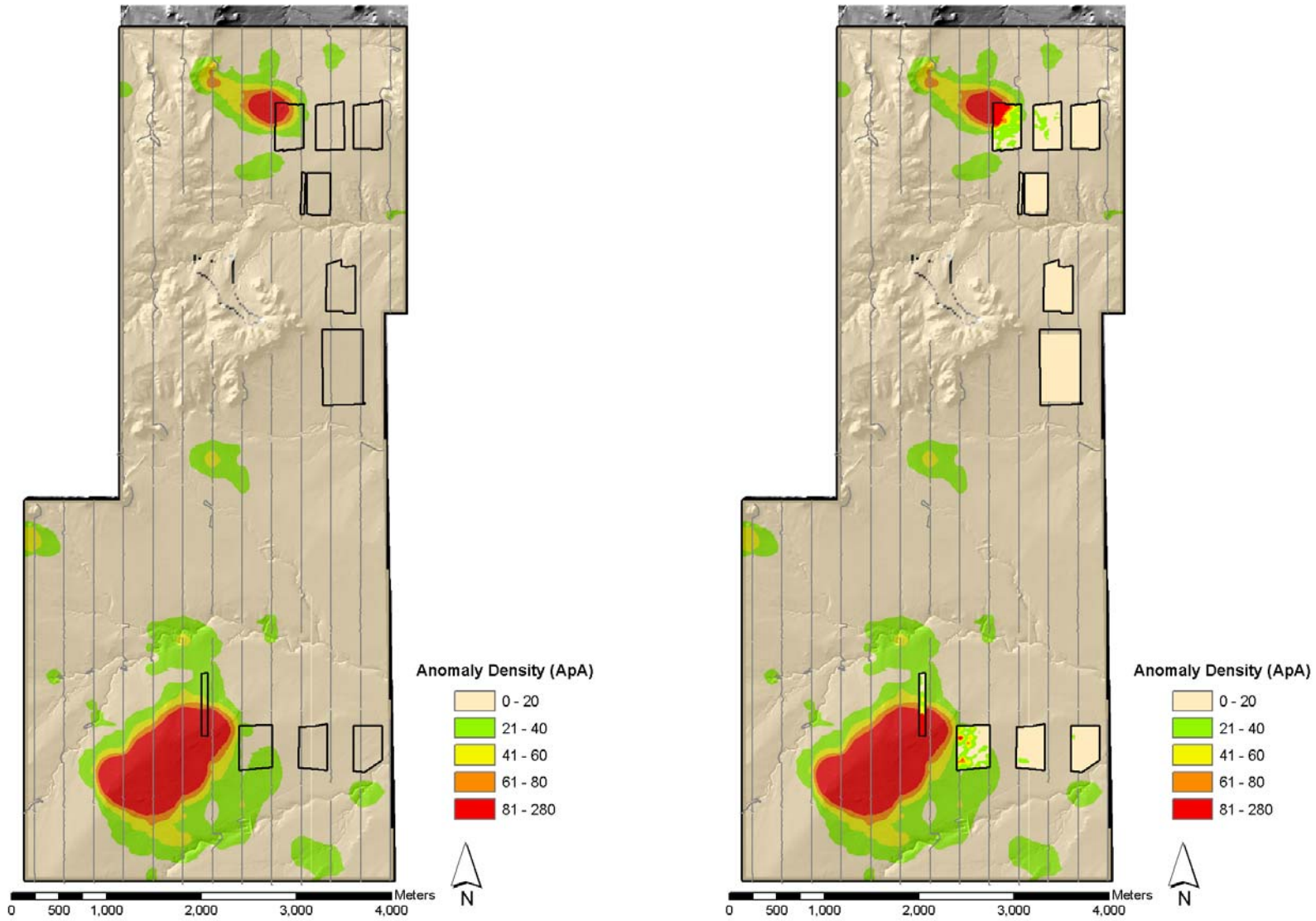


Figure 17. Comparison of kriging estimates of anomaly density developed using sparse transect design data against ground-based full-coverage magnetometer data. Sample transect locations shown as light gray lines. Full-coverage sampling areas shown with heavy black outline. Left image shows only results from kriging estimate; right image shows same kriging estimates with density data from full-coverage areas overlaid using same color scale.

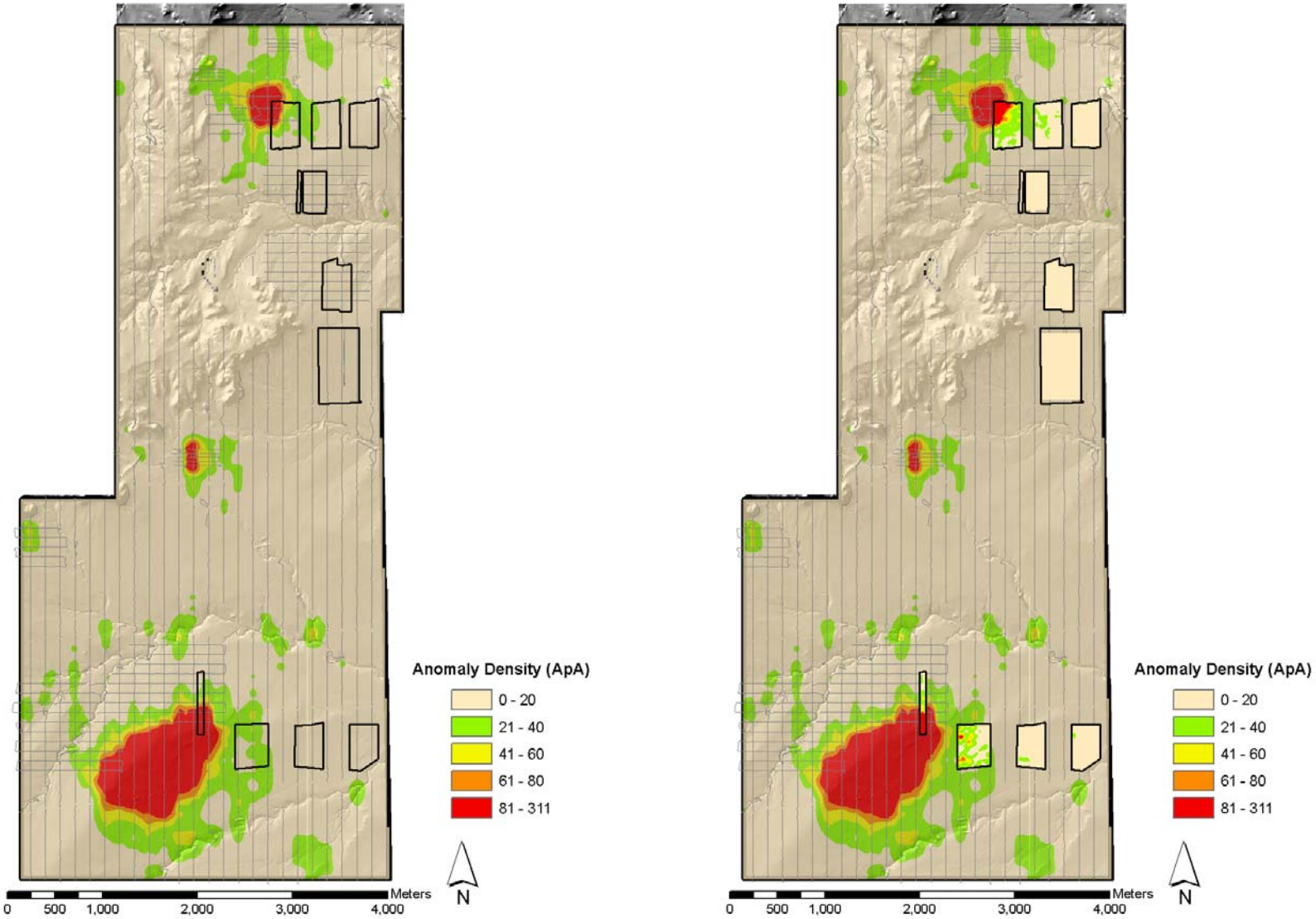


Figure 18. Comparison of kriging estimates of anomaly density developed using dense transect design data against ground-based full-coverage magnetometer data. Sample transect locations shown as light gray lines. Full-coverage sampling areas shown with heavy black outline. Left image shows only results from kriging estimate; right image shows same kriging estimates with density data from full-coverage areas overlaid using same color scale.

4.3.3. High density area delineation comparison

A more specific comparison between the ground-based full-coverage data and the kriging estimates was obtained by applying a density threshold to the full-coverage data to simulate the delineation of a High Anomaly Density Area based on estimated or measured anomaly densities.

The HADA determinations were made for a grid of points covering the extent of the ground-based full-coverage plots and spaced coincident with the grids used for the kriging estimation. This grid definition provided for a point-by-point comparison between the ground-based full-coverage HADA and those determined by the kriging estimates.

In choosing the HADA threshold value to use in comparing the ground-based full-coverage data to the kriging estimate several criteria were considered. Of fundamental concern was that there is a sufficient distribution of locations above the threshold. Because several of the ground-based full-coverage plots have fairly low anomaly density values, lower threshold values were needed. In the end, after examining a range of threshold values, anomaly density threshold values of 10, 20, and 30 ApA were chosen. The use of three different threshold values provided insight into any sensitivity related to the threshold value.

To investigate how well the anomaly density values developed from the transect-based kriging estimates match the HADA identifications made from the ground-based full-coverage data, ROC curves were developed for ground-based full-coverage areas 1 and 3 using each of the three threshold values. The true positive and false positive ratios for the ROC curves were determined using the same methodology as described in Section 4.2.3 of this report.

In order to test the full range of the dense transect kriging estimates, the testing was concentrated on the ground-based full-coverage plots near areas with higher anomaly density. Because of the relatively small size of the ground-based full-coverage plots, the results from several plots were combined to produce the ROC curves. Full-coverage plots 1A, 1B, 1C, and BT4 were combined into one data set for ROC curve analysis, and plots 3A, 3B, and 3C were combined into another data set. A separate ROC curve was developed for each of the two data sets, and for each of the three HADA thresholds listed above. See Figure 8 for the locations of the full-coverage plots.

Figure 19 shows the ROC curves from the Area 1 and Area 3 ground-based full-coverage magnetometer plots for each of the three HADA threshold values. The ROC curves show that the kriging estimates do a good job at representing the spatial and value distribution of magnetic anomaly densities for the full-coverage plots. Defining accuracy as the percentage of correctly identified (positive or negative) cell values and using the ROC threshold level for each plot at the HADA threshold definition, the kriging estimates for the Area 1 plots had an accuracy of 87%, 78%, and 86% for the 10, 20, and 30 ApA HADA definition threshold values respectively. For the Area 3 plots, the accuracy was

75%, 83%, and 90% for the 10, 20, and 30 ApA HADA definition threshold values respectively. These accuracy values are impressive considering the very limited sample data on which the kriging estimates are based and provide additional confirmation that the kriging estimates using limited transect data provide an effective and efficient means of characterizing large former munitions ranges.

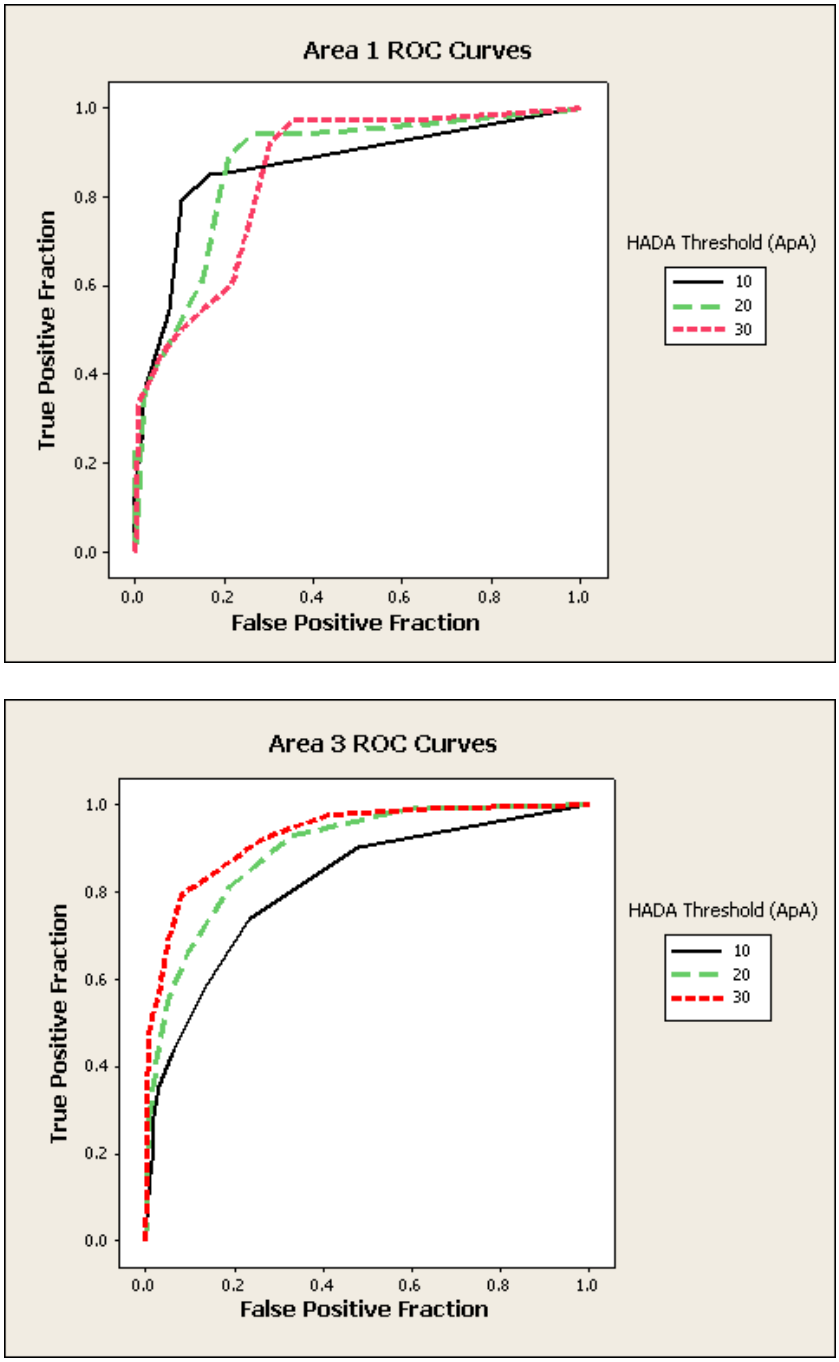


Figure 19. ROC curves for the dense-transect kriging estimates tested against the combined ground-based full-coverage Area 1 plots (top figure) and Area 3 plots (bottom figure).

Figure 20 shows the spatial distribution of the comparison results for the three HADA definition thresholds. In these plots red-colored symbols indicate anomaly density values above the respective threshold; white-colored symbols indicate values below the threshold. The inner symbol shows the result from the dense transect kriging estimate with the outer symbol showing the result from the ground-based full-coverage magnetometer survey. Locations with consistent inner and outer symbol colors show a correct classification from the kriging results. Locations with differing inner and outer symbol colors show a misclassification based on the kriging results as defined in the legend of Figure 19.

The spatial distribution of the classification results shown in Figure 20 show that the classifications are most accurate proximal and distal to the local HADA, and are least accurate at intermediate locations representing transition zones from the HADA to the background (the inferred edges of the target area). This result reflects the sensitivity of the binary classification to small changes in anomaly density value for those areas at or near the HADA threshold.

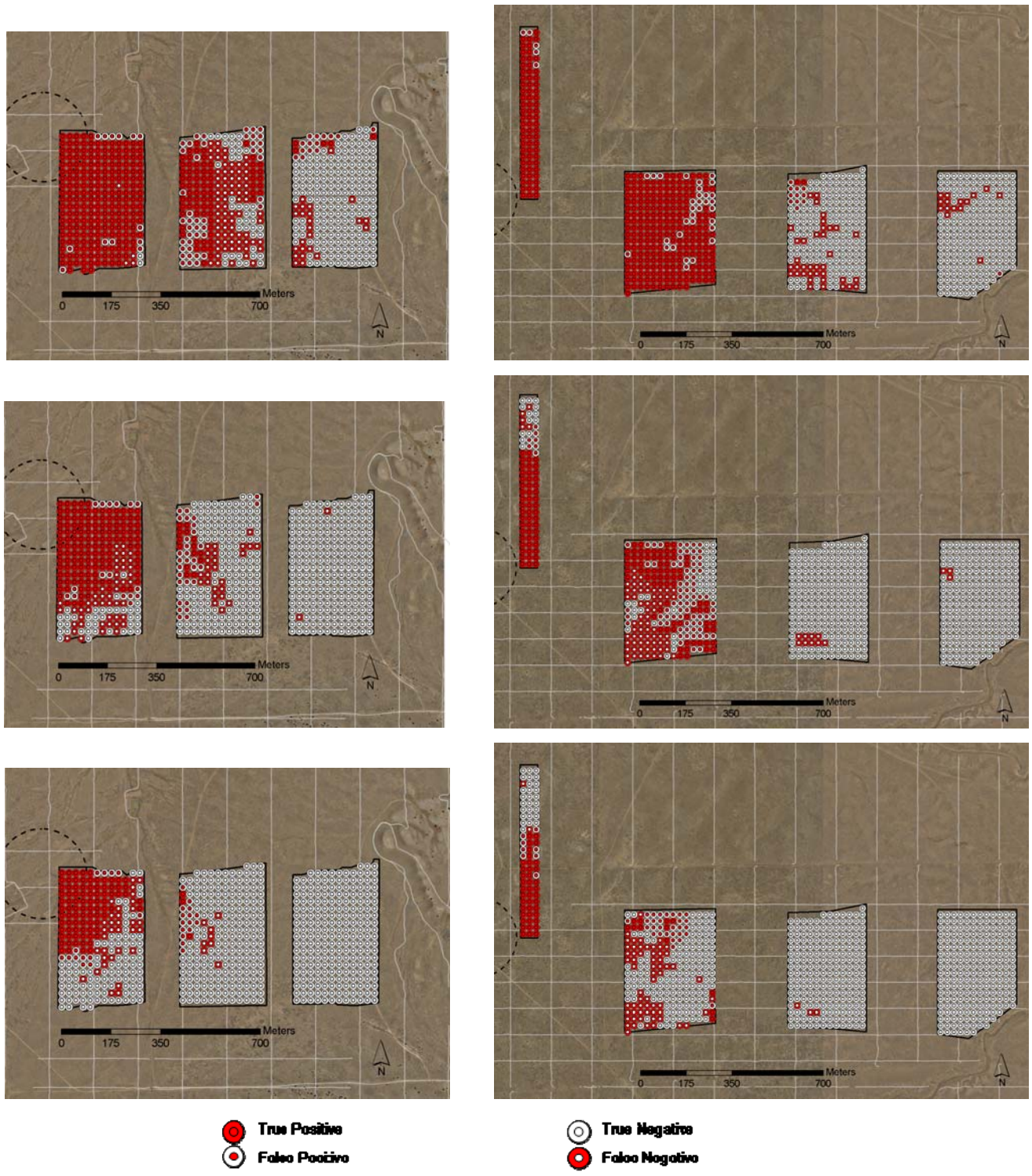


Figure 20. Comparison between kriging estimate of anomaly density (dense transect) and ground-based full-coverage field-measured values for full coverage Area 3 (left column) and full coverage Area 1 (right column). Top, middle, and bottom row represent results at the 10, 20, and 30 ApA HADA threshold levels respectively. Each symbol represents results of the decision made for an individual 25x25 meter grid cell. Inner symbol shows results from kriging estimate; outer symbol shows classification from ground-based full-coverage magnetometer data. Red indicates at or above HADA threshold white indicates below threshold.

5. Transect Sampling Sensitivity Analysis

The nearly complete sampling of the Pueblo WAA site by the helicopter magnetometer system offers an opportunity to test the sensitivity of kriged anomaly density estimates to varying transect configurations. Because the helicopter magnetometer data is continuous across a majority of the site (Figure 4), individual swaths of data designed to simulate transect sampling using a helicopter magnetometer system can be extracted from the data set. Then, using geostatistical estimation techniques, these transect data can be employed to develop spatially continuous estimates of magnetic anomaly density. This approach simulates the process of characterizing a study area using helicopter magnetometer transect data Estimates developed from different transect sampling configurations can be compared against one-another to determine if there are significant differences introduced by the different sampling scenarios. In addition to cross-comparison, the individual estimates can also be compared against the original full-coverage helicopter magnetometer data set to determine the high-density area identification capabilities of the simulated transect data.

The following sections present the construction of different simulated transect sampling data sets using the full-coverage helicopter magnetometry data, the development of kriging estimates of anomaly density from these transect data sets, the identification of high-density areas within the anomaly density estimates, and a comparison between the results generated from the different transect data sets.

5.1. Simulated Transect Data Set Development

Using the original helicopter magnetometer data described in Section 3 of this report (Figure 4), six separate transect sample groups were extracted from the helicopter magnetometer data. Three of the transect data sets were designed with parallel directions, but different grid offset origins; the remaining three were designed with a common grid origin, but differing transect directions. All of the simulated helicopter magnetometer transects had a sampling width of 5 meters and transect spacing between center-lines of 156 meters. The center-line spacing of the simulated helicopter transects (156 meters) is the same as that used for the north-south transects of the ground-based dense transect magnetometer survey (Figure 6).

The first sample transect data set, labeled as NS, was designed to follow the same north-south transect layout as used in the original ground-based magnetometer dense transect sampling (Figure 6). In this scheme, parallel transects were laid out in a north-south direction with approximately 156 meters spacing between the transect center-lines. Because of the elongate north-south configuration of the PPBR study site, the NS data set represents the most probable configuration if planning helicopter-based magnetometer transect sampling for the Pueblo WAA site. No orthogonal transects were used in this testing. Two additional north-south transect data sets were also constructed from the helicopter magnetometer data. Data sets NS52E and NS52W represent transects parallel

to the NS data set, but offset by 52 meters east and 52 meters west respectively. The 52 meter offset represents 1/3 of the 156 meter transect spacing. Offsetting the transect origin by 1/3 the transect spacing allows three parallel but independent data sets to be constructed. These data sets were developed to determine if using different transect origins would have any significant impact on the subsequent geostatistical estimates.

Three additional transect data sets were developed from the full-coverage helicopter magnetometer data to test the impact of sample transect orientation on the geostatistical estimates. These data sets used the same 156 meter spacing as the previous sampling transects, but differing orientations. The first data set, labeled EW, was oriented 90° from the NS data set, in an east-west configuration. The remaining two data sets were oriented in a southwest-to-northeast and northwest-to-southeast orientation and were labeled as SWNE and NWSE respectively. Figure 21 shows the sample transect locations and identified anomalies used for each of the six data sets.

The six sample transect groups extracted from the full-coverage helicopter magnetometer data represent raw anomaly locations as would be provided from a hypothetical airborne magnetometer survey. These data consist of the spatial location of each detected magnetic anomaly. To use these data to provide anomaly density estimates, the raw anomaly location data must first be transformed to spatially-varying density values. This was accomplished by overlaying a grid with the same origin and grid-spacing as that used to represent the gridded helicopter magnetometry density values (Figure 10) onto the sample transect data, and averaging the anomaly counts within a moving window for each grid cell location. A spatial density value was then computed by dividing by the transect area falling within the averaging window. Density values were only computed for grid cells falling within the actual survey transect footprint. It is this anomaly density data which is presented below and used as input data for subsequent kriging analyses.

Table 8 lists summary information regarding the transect anomaly density data sets computed from the six simulated helicopter transect data sets. As shown in this table, although the maximum density from the simulated transects is well below the actual maximum, the other population parameters are well represented by the transect samples. Figure 22 depicts the transect anomaly density distribution for each of the six data sets as analyzed in Visual Sample Plan (VSP) (Matzke, et al., 2007). Each histogram shows the log counts of the densities calculated from a 300 meter diameter window. For a more detailed description of the window density calculations see ESTCP (2006a).

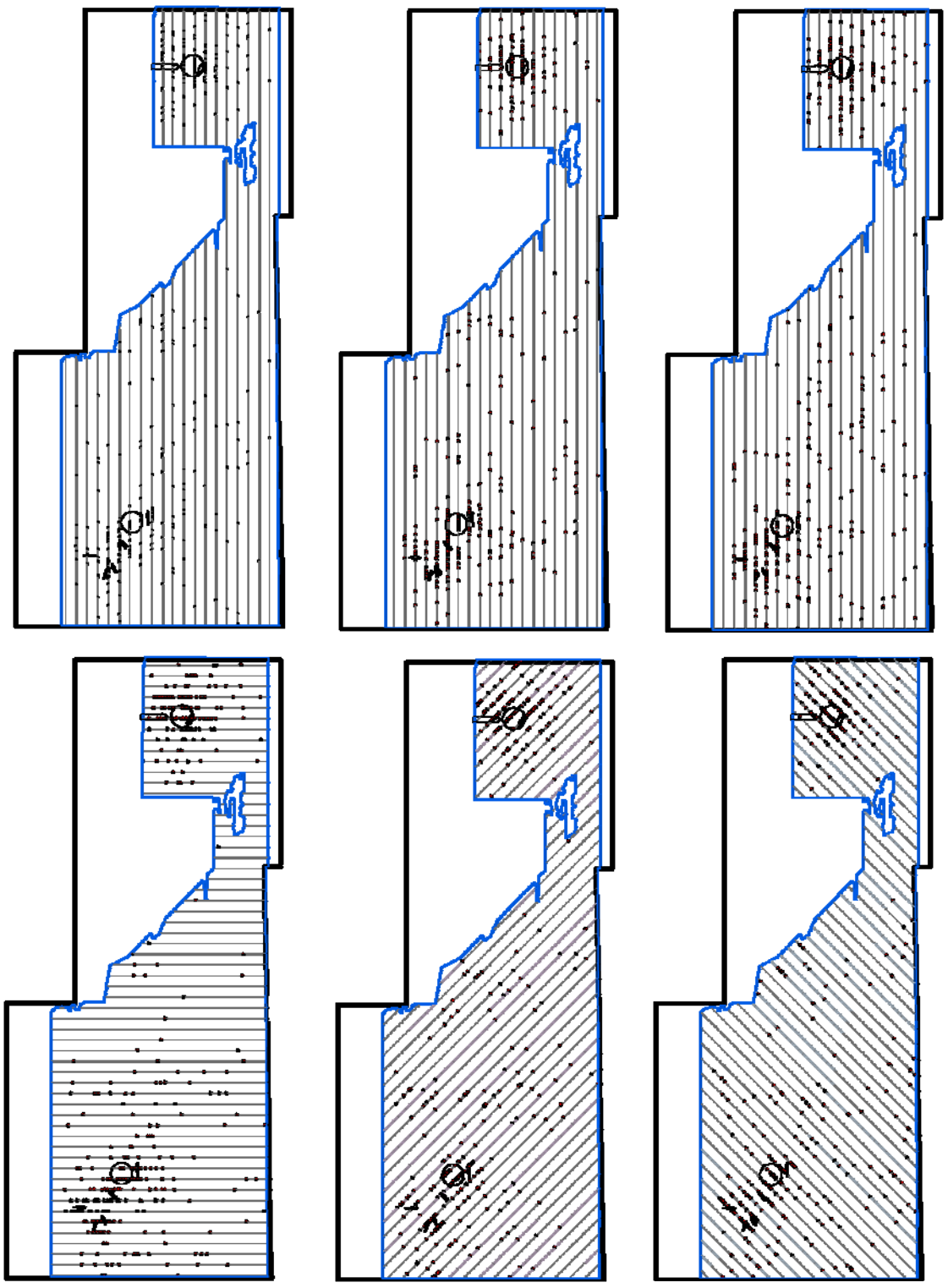


Figure 21. Sample transect configurations used with helicopter magnetometry data. Top row shows (from left to right) transect locations for the NS, NS52E, and NS52W data sets. Bottom row shows (from left to right) transect locations for the EW, SWNE, and NWSE data sets. Blue outline shows extent of helicopter magnetometry data; individual magnetic anomalies shown as points. Circular target locations shown for reference.

Table 8. Summary information for the six simulated helicopter transect data sets. The Area Sampled column lists the percentage of the helicopter magnetometer sample area covered by the transects; Total Length column lists combined length of transects. The average (Mean), standard deviation (SD), maximum (Max.), median, minimum (Min.), and various percentiles for the data set anomaly density values are also listed. Last row (FULL) presents information from the original full-coverage helicopter magnetometer data set.

Data Set	Area Sampled	Total Length (km)	Density (anomalies per acre)								
			Mean	SD	Max.	95th %	75th %	Median	25th %	5th %	Min.
NS	3.1%	126.8	2.5	5.3	47.3	12.5	2.5	0	0	0	0
NS52E	4.0%	160.8	2.4	6.1	62.3	12.5	2.5	0	0	0	0
NS52W	4.1%	164.5	2.7	5.5	44.8	12.5	2.5	0	0	0	0
EW	3.9%	158.1	2.5	5.6	64.8	12.5	2.5	0	0	0	0
SWNE	3.2%	129.8	2.6	6.2	63.4	13.4	3.4	0	0	0	0
NWSE	3.9%	156.3	2.8	6.8	64.9	13.4	3.4	0	0	0	0
FULL	100%	*NA*	2.5	6.4	106.4	12.0	2.4	0.1	0	0	0

As shown in Table 8, while only sampling approximately 3-4% by area of the site, the transect sampling provides a good representation of the mean anomaly density of the helicopter magnetometry data. The data in Table 8, and depicted in Figure 22, show no clear bias based on sample transect design indicating that each of the transect data sets represent the underlying population of helicopter magnetometer density values equally well. Because each of the histograms in Figure 22 is very similar, it is likely that the selected critical density for each design would be about the same. However, the final selected critical density would also incorporate the spatial location of the identified high density areas from each design.

The histograms from Figure 22 and spatial density maps using different density thresholds could be used in a separate analysis on each of the designs shown in Figure 21. These analyses would result in finalized high density area boundaries which could be used in a comparison. Instead of going through the steps of a typical analysis for each design, Sections 5.2 and 5.3 present a methodology to compare all six transect data sets to the high density areas identified from the helicopter full-coverage using a single threshold density.

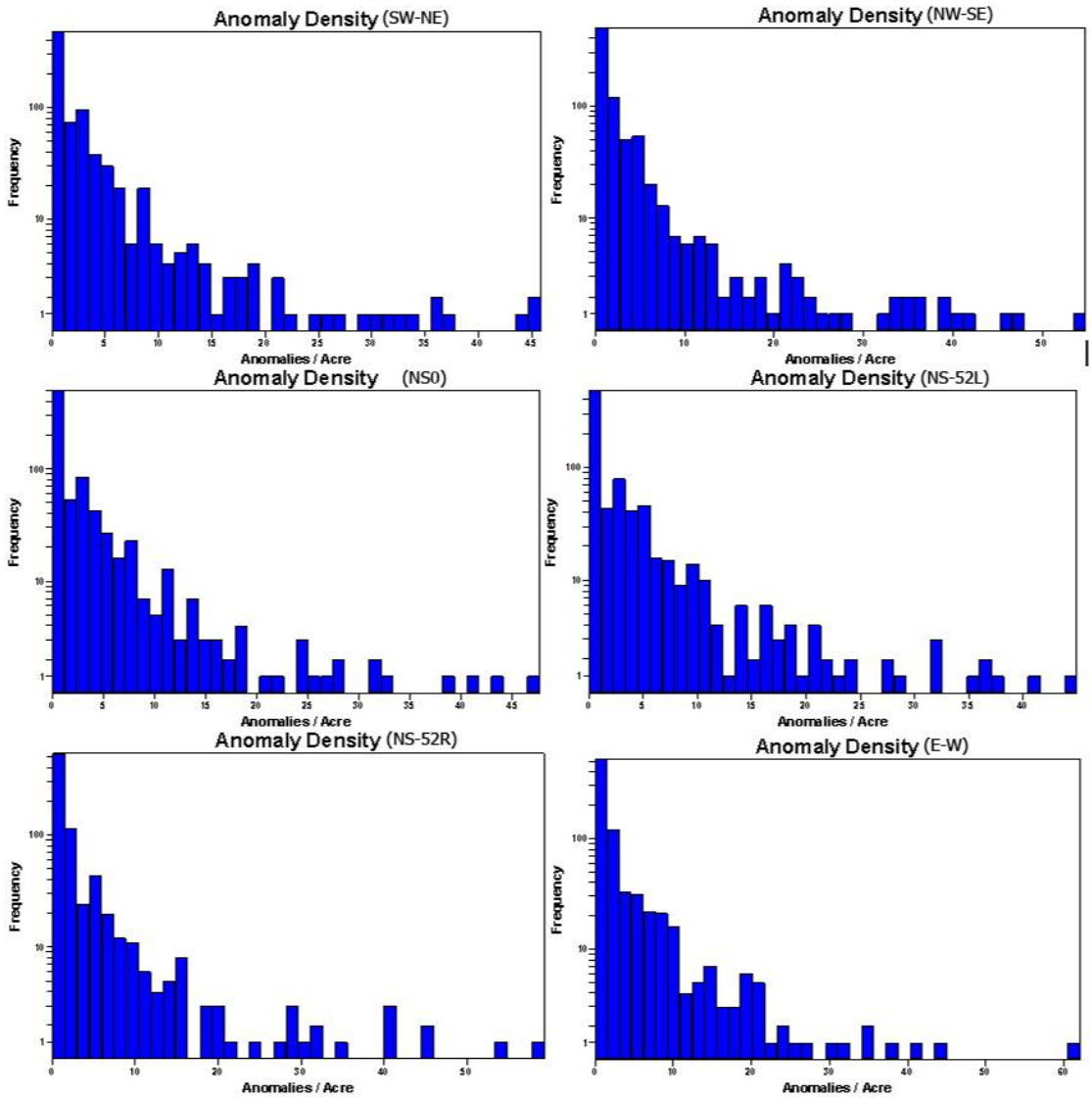


Figure 22. Histograms of 300 meter diameter window densities for each of the six different 156 meter spaced simulated helicopter transect placements.

5.2. Kriging Density Calculations and High-Density Area Threshold Selection

High Anomaly Density Areas (HADA) are often indicative of former range target areas and so are important in the characterization of these sites. Analyses were performed to test the ability of the six transect designs discussed in Section 5.1 to accurately locate and delineate HADA. For this, a kriging estimate of anomaly density was developed from each of the six transect data sets extracted from the full-coverage helicopter data. Each kriging estimate was developed independently using individual data sets and semivariogram models. These kriging estimates of anomaly density were then used to identify HADA based on a specific anomaly density threshold. These areas were then compared to HADA identified in the full-coverage helicopter magnetometer data set using the same threshold value.

In order to test the HADA identification capability of the six transect data sets, some defined density threshold must be established to delineate these areas. A HADA threshold of 12 ApA was used in this analysis. This is the same threshold value as used in the Section 4.2.3 of this report.

Figure 23 shows the HADA defined by the 12 ApA threshold using the quadratic kernel smoothed, full-coverage helicopter magnetometer data. These are the locations which will be used in testing HADA identification from the kriged estimates of anomaly density created using the different configurations of helicopter magnetometer transect data.

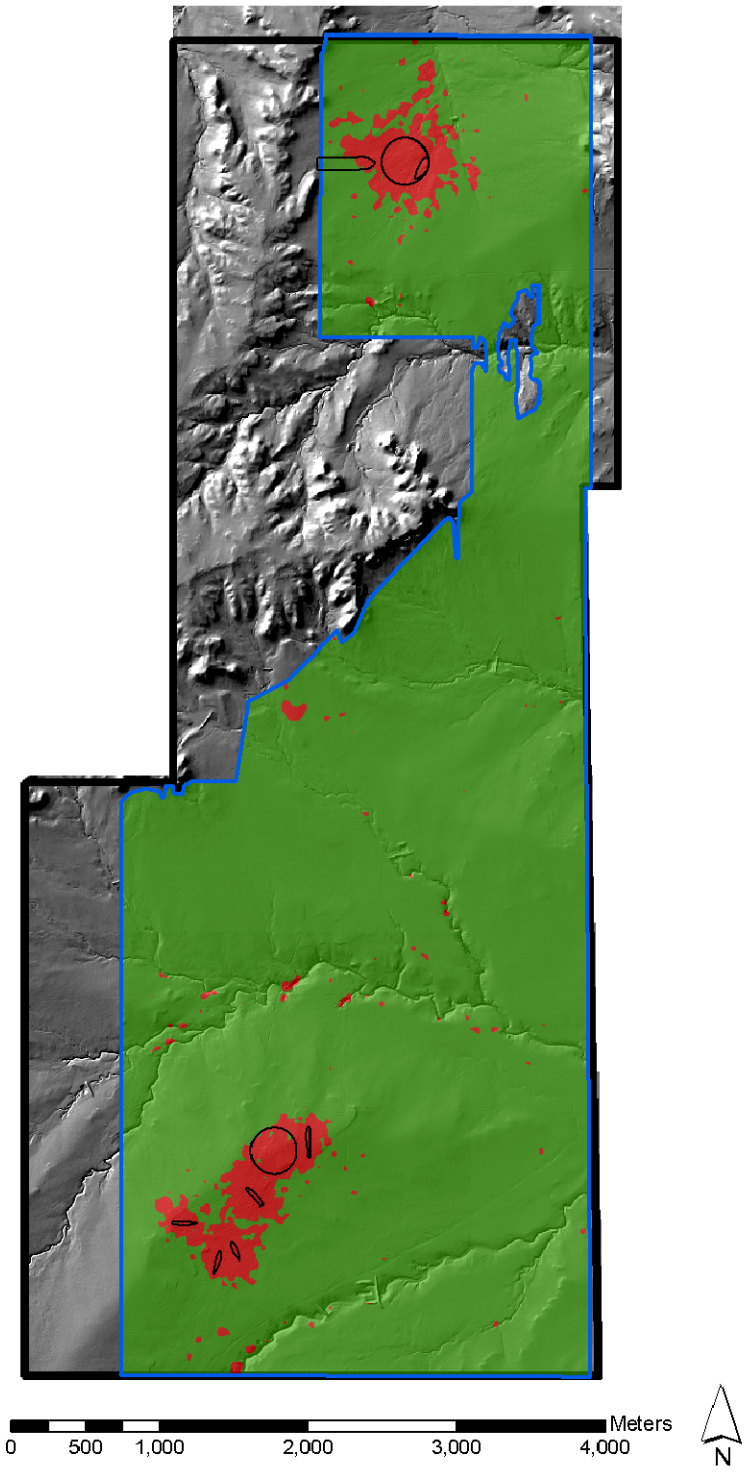


Figure 23. Distribution of HADA in the full-coverage helicopter magnetometry data. HADA defined as grid cells with anomaly densities at or above 12 ApA. Areas at or above threshold shown in red shading; areas below threshold shown in green shading. Target circles and ship outlines shown for reference.

Simulated helicopter magnetometer transects were extracted from the full-coverage helicopter data set using the six sampling configurations shown in Figure 21. These transects were then used to develop six separate estimates of anomaly density using established geostatistical techniques (Ordinary Kriging). The kriged estimates from these six data sets are shown in Figure 24. In this figure red shaded areas show locations with anomaly density values at or above the 12 ApA HADA delineation threshold. As shown in Figure 24, the general spatial distribution of locations at or above 12 ApA is similar for each of the kriged density estimates. Each of the six independent estimates identifies large target masses in the northern and southern portions of the study area. This result provides confidence that each of the general HADA at the PPBR study site as identified in the full-coverage helicopter magnetometer data would have been identified using helicopter magnetometer transect sampling with a 156 m spacing regardless of the orientation or origin of the transect sampling scheme. Identification of these HADA from very limited sampling is the main goal of this work. For a site that is amenable to airborne surveys, once a HADA is identified from the transect sampling, the helicopter could be tasked to fill in the details of the HADA boundary and the density variation within the HADA.

The complete range of magnetic anomaly values estimated from each of the six transect sampling schemes is shown in Figure 25. In this figure, the anomaly density values are shown by color-shading. This figure shows how many of the target feature details are maintained in the kriging estimate even though only a small portion of the site was covered by transect sampling (compare to Figure 10). Although the helicopter transect kriging estimates shown in Figure 25 readily identify the major HADA locations, they do not contain as much detail regarding the anomaly density distribution as do estimates created using ground-based transect sampling (Figure 17 and Figure 18). This result is due to the filtering aspect of helicopter-based surveying. This fact should be considered when considering different technologies for transect sampling.

The total acreage represented by the different simulated helicopter transect surveys listed in Table 8 ranges between 150 and 200 acres. Because of the rapid collection rate of the helicopter magnetometry system, averaging about 400 acres per day (ESTCP, 2006b), it would be possible to create any of the maps shown in Figure 25 in 1 to 2 days, including both data acquisition and processing time. The use of helicopter transect sampling for initial site investigation could allow for very rapid characterization of large sites adding additional efficiency to the Wide Area Assessment process.

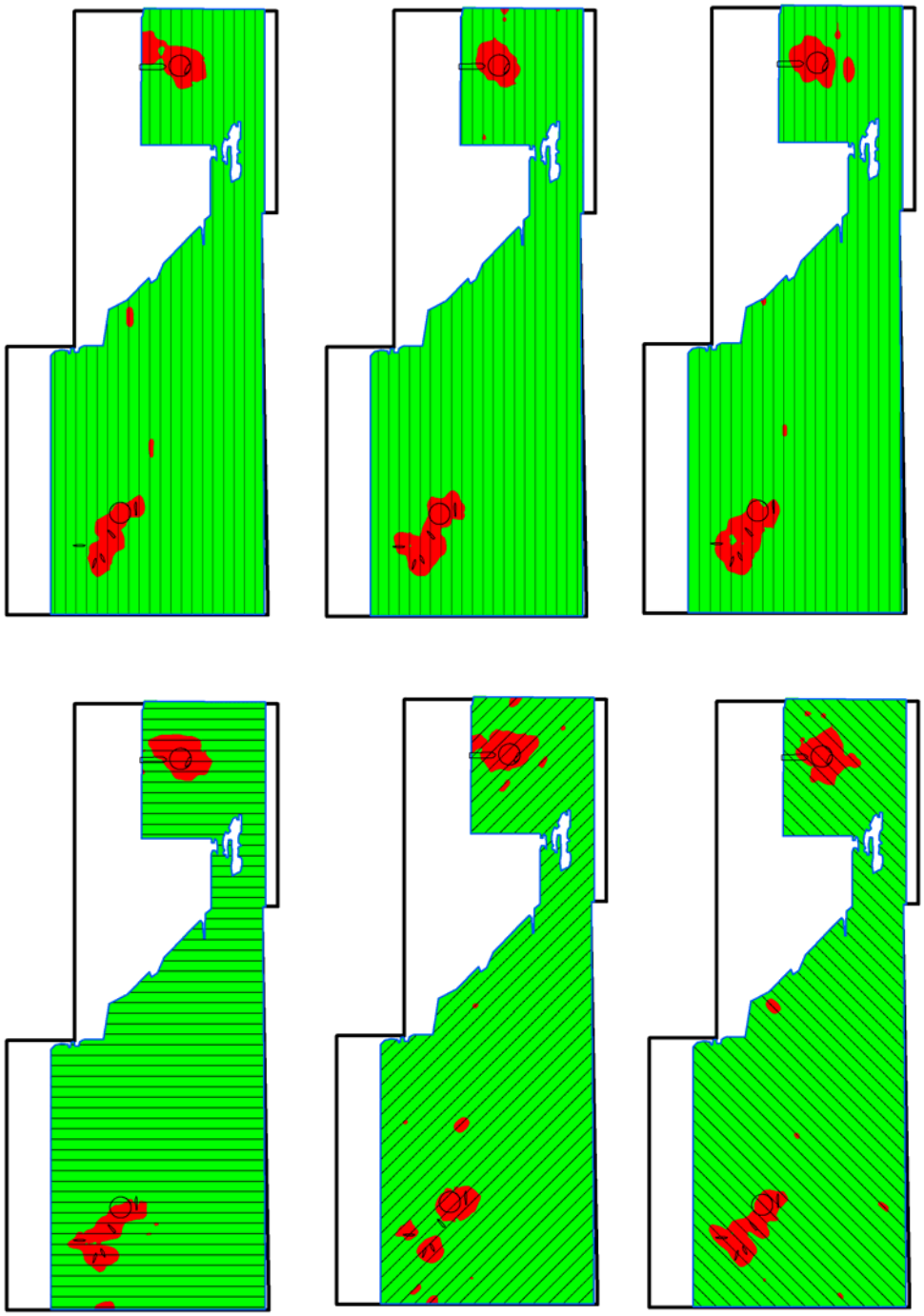


Figure 24. Distribution of HADA in kriging density estimates. HADA are defined as grid cells with anomaly densities at or above 12 ApA. Areas at or above threshold shown in red shading; areas below threshold shown in green shading. Sample transect configurations shown as black lines. Top row shows (from left to right) locations above threshold for the NS, NS52E, and NS52W data sets. Bottom row shows (from left to right) locations for the EW, SWNE, and NWSE data sets. Blue outline shows extent of helicopter magnetometry data. Circular target locations and ship outlines shown for reference.

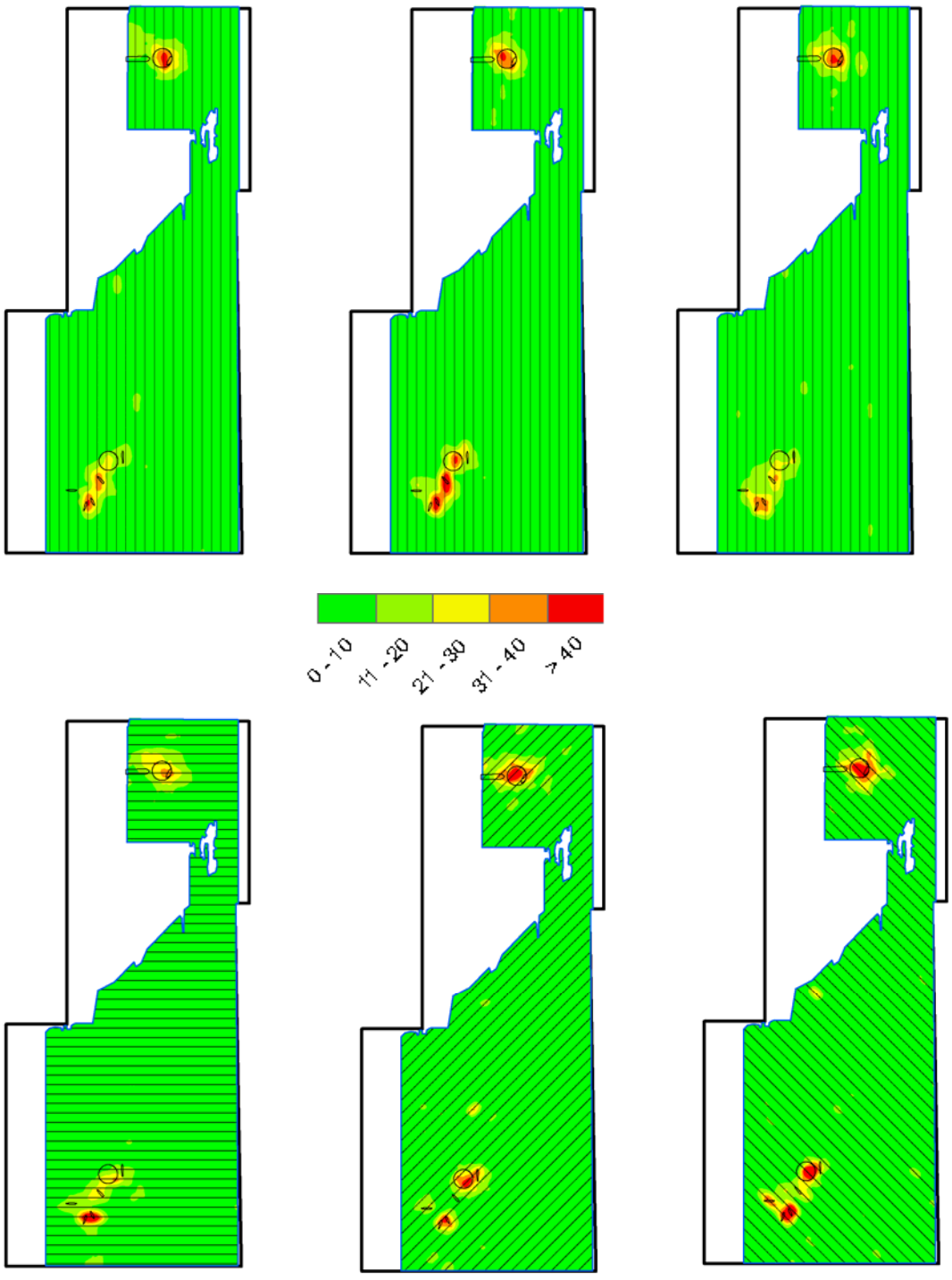


Figure 25. Kriged estimates of anomaly density (ApA) developed from the six transect sampling scenarios. Sample transect configurations shown as black lines. Top row shows (from left to right) kriged estimate of anomaly density for the NS, NS52E, and NS52W data sets. Bottom row shows (from left to right) density estimates for the EW, SWNE, and NWSE data sets. Blue outline shows extent of helicopter magnetometry data. Circular target locations and ship outlines shown for reference.

5.3. High-Density Area Identification and Delineation Comparison

To investigate how well the anomaly density estimates developed from the simulated transect sample data match the HADA identifications made from the full-coverage helicopter data, ROC curves were developed for each of the six kriging estimates. The ROC curves show the relationship between true positive and false positive ratios for various threshold levels. For this analysis, true values are defined as those grid cells in the full-coverage helicopter data with an anomaly density at or above 12 ApA (Figure 23). These correspond to our previously defined HADA. These will be the basis to compare the results from the kriging estimates to determine the false positive and true positive ratios. The true positive and false positive ratios for the ROC curves were determined using the same methodology as described in Section 4.2.3 of this report.

Figure 26 shows the ROC curve developed using the kriging estimates of anomaly density based on the NS transect data set. As shown by this figure, the ROC curve plots in the extreme upper left-hand corner of the graph (e.g., achieving 90% true positive ratio with approximately 15% false positive ratio) indicating that this model is effective and efficient in identifying HADA in the helicopter magnetometer data. The steep rise and sharp bend in the initial portion of the curve indicates that the model (kriged estimate of transect anomaly density) is good at identifying the same high-density areas as the full-coverage data set (high sensitivity) without misidentifying HADA outside of the true HADA areas (good specificity).

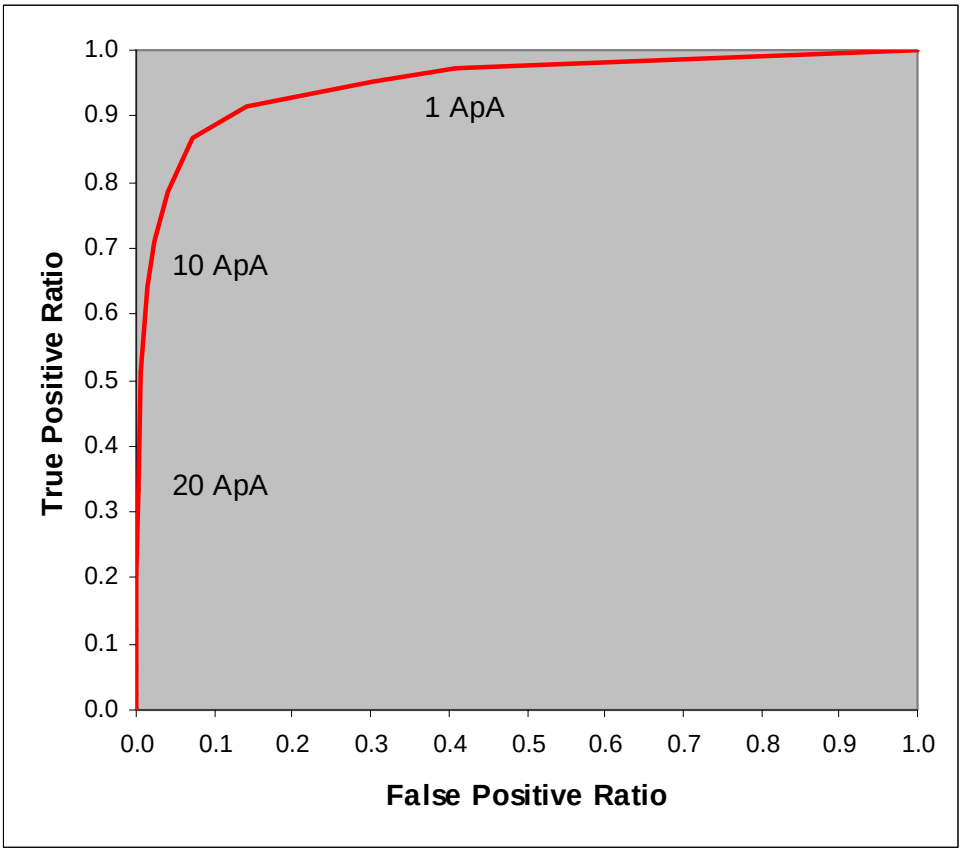


Figure 26. ROC curve (red line) for HADA identification using kriging estimate of anomaly density based on NS transect data set. Ratio values were computed using a series of density threshold values applied to the kriging estimates ranging from zero to the maximum estimated value. Example threshold values along the ROC curve are shown by "ApA" labeled values.

ROC curves of HADA identification capability were developed for each of the six transect sampling scenarios included in the sensitivity analysis. The ROC curves from these six scenarios are shown in Figure 27. As this figure shows, all the ROC curves from the six individual estimates of high-density grid locations cluster tightly together. This indicates that the high-density area identification performance is not greatly affected by the orientation or origin of the transect sampling design. The general location of the curves (upper left corner of plot space) also indicates that all the kriging estimates are effective in properly identifying the same HADA as identified in the full-coverage data set.

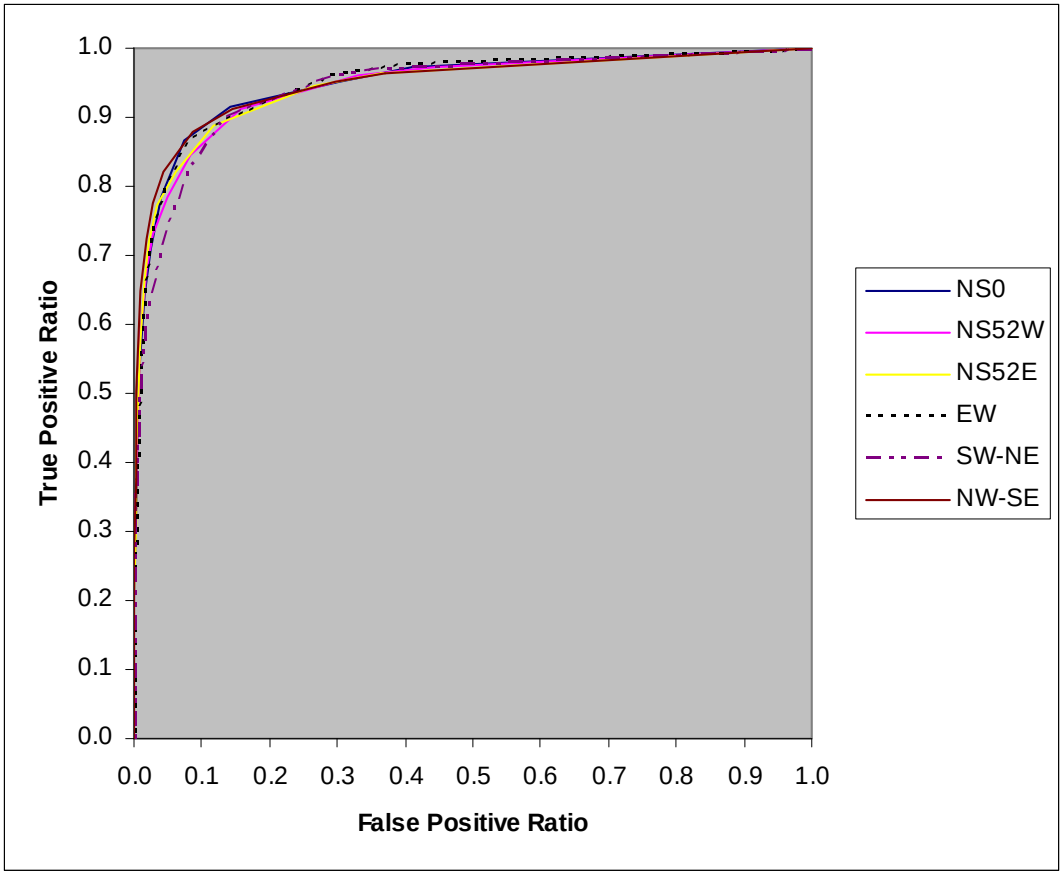


Figure 27. ROC curves for HADA identification using kriging estimate of anomaly density based on simulated transect sample data sets. Ratio values computed using a series of density threshold values applied to the kriging estimates ranging from zero to the maximum estimated value.

The area under a ROC curve is indicative of the predictive performance of the model being tested. A model with perfect performance would have an area of 1; a model with a ROC curve falling on the diagonal (no better than random assignment of target areas) would have an area of 0.5. Models with useful predictive capabilities have ROC curve areas between 0.5 and 1.0, with the more effective and efficient models having areas closer to 1.0.

Table 9 lists the area under the ROC curve for the HADA location predictions from each of the six transect sampling scenarios used in the sensitivity analysis. The small range in area values (0.86 to 0.92) confirms the observation from Figure 27, that the six different kriging estimates of anomaly density have similar predictive capabilities for locating high anomaly density areas.

Table 9. Area under ROC curve for each prediction of HADA developed from the six transect sampling scenarios.

Transect Data Set	ROC Curve Area
NS	0.88
NS52W	0.90
NS52E	0.89
EW	0.86
SWNE	0.88
NWSE	0.92

The goal of developing the six different transect sampling schemes was to determine if the orientation or origin of the transects have a significant impact on the identification of high anomaly density areas. To test the sensitivity of HADA identification performance, the ROC curves from the six transect designs will be compared to one-another using the Wilcoxon statistic. The Wilcoxon statistic provides a quantitative comparison of variables between two sample groups to determine if they are from the same population. Applying this test to the results from the six transect sampling scenarios will determine if there are significant differences between these transect designs. Significant differences between the transect designs would indicate a sensitivity to transect design parameters (orientation or origin).

Using the techniques presented in Fogarty et al. (2005), the Wilcoxon statistic was computed for each unique pairing of ROC curve areas from the six different transect sampling scenarios. This pairing tested each ROC curve against each of the other five curves, resulting in a total of 15 unique comparisons. The null hypothesis (H_0) for the Wilcoxon test is that there is no significant difference between the areas beneath the two ROC curves under consideration. Rejection of H_0 indicates that there is a significant difference between the two ROC curve areas. Failing to reject H_0 indicates that there is not a significant difference between the ROC curve areas.

For each of the 15 comparisons, the standard error for the area under the ROC curves was computed and used to compute a Z value. The probability of obtaining that Z value (p-value) was then determined and used to identify significantly different ROC curve areas. The p-value represents the probability of obtaining a specific Z value or greater, assuming H_0 (no difference) is true. Small p-values indicate a small probability that the observed ROC curves have the same area under the curve and would lead to rejection of the null hypothesis, indicating a significant difference between the areas of the ROC curves. An overall significance level of 0.05 (α) was desired for the analysis. To guard against the effects of multiple testing on the overall α, a Bonferroni correction was applied. This correction divides the overall α by the number of comparisons (c) to identify the conservative α^* $\left(\alpha^* = \alpha / c\right)$. For this comparison, $\alpha^* = .05/15 = .00333$ was used. P-values below this level indicate a significant difference between two ROC curve areas.

Table 10 lists the p-values for each of the 15 unique pairing combinations of the six transect sampling scenarios. The upper and lower portions of the tables are symmetric reflecting the same comparisons listed in different order. Any probability value below α^* would lead to a rejection of the null hypothesis indicating significant differences between the ROC curve areas.

Table 10. P-values for each of the pair wise comparison tests.

	NWSE	SWNE	NS52E	NS52W	EW	NS
NWSE		0.224	0.383	0.594	0.062	0.307
SWNE	0.224		0.730	0.581	0.514	0.919
NS52E	0.383	0.730		0.810	0.318	0.826
NS52W	0.594	0.581	0.810		0.252	0.670
EW	0.062	0.514	0.318	0.252		0.476
NS	0.307	0.919	0.826	0.670	0.476	

An examination of Table 10 indicates that none of the p-values are less than α^* (0.00333). This demonstrates that, for the data sets investigated here, there is no significant difference in ROC curve area based on differences in transect orientation or in transect origin. Each of the transect configurations investigated here perform equally well in identifying HADA, based on a ROC curve area criterion, using the test configuration and data sets presented here.

The above analysis provides a test of some of the underlying assumptions inherent in the design of transect sampling for WAA sites. Specifically, these analyses address the assumption that the neither the orientation nor the origin of the transects is significant when trying to locate HADA assuming the transect spacing is appropriate and the underlying anomaly density pattern has no preferred spatial orientation (O'Brien et al., 2005). The results from the analysis show that, for the data sets examined here, there is no significant sensitivity to the orientation or origin of the transect sampling scheme. This conclusion further supports the use of transect sampling in the WAA process.

6. Summary

The purpose of this report was to compare geostatistical estimates of anomaly densities developed from statistically based transect surveys covering 1-3% of the site area, to anomaly density values derived from air and land based full-coverage magnetometer surveys. These comparisons provide useful information regarding the effectiveness of using survey transect data in site characterization. As part of these comparisons, ratios comparing anomaly counts recorded by the helicopter and ground-based magnetometer systems were developed for a number of subregions of the site. In addition, several alternate transect sampling data sets were constructed from the full-coverage helicopter data. These data sets were used to investigate the sensitivity of target area delineation to several transect design parameters.

The main focus of this report was to evaluate kriging estimates of anomaly density developed using magnetometer transect data against full-coverage (continuous sampling) magnetometer data sets collected at the same site. Although direct comparisons of the helicopter magnetometry to the kriging estimates of anomaly density were restricted due to the differences in sensor stand-off between the ground-based and helicopter-based systems, ROC curve results indicate that the kriging estimates readily identified the HADA observed in the full-coverage helicopter data. Analysis of anomaly counts where both ground-based and helicopter-based full-coverage data were available reveals helicopter/ground-based anomaly count ratios of between 0.29 and 0.08. These results provide an expected range between anomaly counts determined using the helicopter-based and ground-based magnetometer systems employed at the PPBR study site. Because the kriging estimates of anomaly density were based on ground-based magnetometer transect data, the relative value of the helicopter/kriging anomaly counts is informative in relating the kriging estimates to the helicopter magnetometry. If the kriging estimates of anomaly density are to be considered representative of that which would be obtained from a ground-based, full-coverage survey of the entire site, then the ratio of helicopter and kriging anomaly counts should fall within the helicopter/ground-based anomaly count range (0.29 – 0.08). The helicopter/kriging ratios for the sparse and dense transect scenarios were 0.14 and 0.13 respectively. These values fall within the range of values for the helicopter/ground-based full-coverage sites and they provide confidence that the geostatistical estimates of anomaly density are comparable to those that would result from a ground-based full-coverage magnetometer survey of the site.

Although covering a small area compared to the helicopter-based full coverage data set, several ground-based full coverage magnetometer data sets were available for portions of the PPBR study site. A more direct comparison of the data from these data sets was possible since the same sensor platform was used in these data sets as was used to collect the transect data used to develop the kriging estimates. These comparisons show good general agreement between the kriged estimates and the measured values. Total anomaly counts for the ground-based full-coverage areas are 4,059 and 4,267 for the sparse and dense transect kriging estimates respectively, compared to 5,619 from the measured data. The estimated values have impressive accuracy considering they are based on sampling

only 0.7% and 2.0% of the full-coverage sites respectively. Site by site anomaly density comparisons (Figure 13 through Figure 16) show that the general distribution of anomaly densities, as determined by the ground-based full-coverage surveys, is well represented in the kriging estimates. The exception to this is that the kriging estimates, in general, do not provide a good estimate of the maximum anomaly density. This is to be expected from estimates based on limited transect sampling because of the small probability of encountering the maximum population value in the limited sampling. Comparisons of the spatial pattern of anomaly densities show that the spatial patterns in the kriging estimates compare favorably with those from the ground-based full coverage data sets (Figure 17 and Figure 18).

The full-coverage helicopter magnetometer data set provided the opportunity to test the effect of different transect sampling scenarios on kriging estimates of anomaly density developed from these transects. This test was implemented by developing six different simulated transect data sets by extracting anomaly location information from the full-coverage helicopter data. Each transect data set was designed using a different origin or transect orientation. Kriging estimates developed from each of the transect data sets were then compared to the original full-coverage data. ROC curve analysis and graphical comparisons indicate that the simulated transects provide adequate information to identify high anomaly density areas in the kriging estimates. The ROC curve results demonstrate that the application of geostatistical estimation (kriging) to the transect data provides a highly effective means of identifying high anomaly density locations at the PPBR site. In addition, statistical testing of the ROC curves from the different transect scenarios indicates that there is no significant difference between them. This supports the idea that at sites without strong asymmetry in the underlying anomaly density pattern, for the same transect spacing, the orientation and origin of the transect sampling plan will not significantly affect the results of the kriging estimation.

Based on the analyses presented above, if an appropriate transect design is developed (per the VSP transect sampling design module), the process of identifying and accurately delineating high anomaly density areas from kriging estimates developed from relatively sparse transect survey data is very successful. In addition, the magnitudes and spatial patterns of anomaly density in the estimates compare very well against full-coverage survey data, demonstrating that this process is well suited to the task of site characterization of large former munitions ranges. In addition, the results from the simulated helicopter transect surveys show that this process can be equally well applied using helicopter magnetometer survey data. This raises the possibility of implementing Wide Area Assessment surveys using a helicopter magnetometer system to collect transect data. Based on the results in this report and the rapid collection rate of the helicopter magnetometry system, it should be possible to conduct a helicopter transect survey and produce a kriged map of anomaly density values for the entire Pueblo WAA site, in as little as 1 to 2 days. It has been shown here that maps created from this type of survey would be effective in identifying potential target locations. This initial survey could be used to focus additional survey work to better define the boundary of potential target areas or to reduce uncertainty in suspect areas. This approach could lead to significant improvements in the efficiency of characterizing large munitions ranges.

References

ESTCP, 2006a, Application of Statistically-Based Site Characterization Tools to the Pueblo Precision Bombing and Pattern Gunnery Range #2 for the ESTCP Wide Area Assessment Demonstration, ESTCP Project # 200325, Draft #1, Environmental, Security Technology Certification Program.

ESTCP, 2006b, Wide Area Assessment Interim Report for Phases I and II, LiDAR, Orthophotography, and Helicopter MTADS at Pueblo Precision Bombing Range #2, ESTCP Project # 200416 and 200535, Environmental, Security Technology Certification Program.

Fogarty, J., Baker R.S., and Hudson, S.E., 2005, Case Studies in the use of ROC Curve Analysis for Sensor-Based Estimates in Human Computer Interaction, ACM International Conference Proceedings Series, v. 112, p. 129-136.

Harbaugh, G.R., Steinhurst, D.A., and Khadr, N., 2006, Wide Area UXO Contamination Evaluation by Transect Magnetometer Surveys, Pueblo Precision Bombing and Pattern Gunnery Range #2 Demonstration Data Report, NOVA-2031-TR-0001, NOVA Research, Inc., Alexandria, VA.

Isaaks, E.H. and R.M. Srivastava, 1989, An Introduction to Applied Geostatistics, Oxford University Press, Inc., New York, NY.

Matzke, B.D., Wilson, J.E., Nuffer, L.L., Dowson, S.T., Gilbert, R.O., Hassig, N.L., Hathaway, J.E., Murray, C.J., Sego, L.H., Pulsipher, B.A., Roberts, B., and McKenna, A., 2007, Visual Sample Plan Version 5.0 User's Guide, Pacific Northwest National Laboratory publication PNNL-16939.

O'Brien RF, Carlson DK, Gilbert RO, Wilson JE, Bates DJ, Pulsipher, BA (2005) Statistical algorithms for designing geophysical surveys to detect UXO target areas. Subsurface Sensing Technologies and Applications 6, 251–270.

Schabenberger, O., and Gotway, C.A., 2005, Statistical Methods for Spatial Data Analysis, Chapman & Hall/CRC, Boca Raton, Fl.

Versar, 2007, Pueblo Precision Bombing and Pattern Gunnery Range #2, Conceptual Site Model, Draft Version 3.1, Versar, Inc., Springfield, Va.